

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
3 October 2002 (03.10.2002)

PCT

(10) International Publication Number  
**WO 02/077213 A2**

- (51) International Patent Classification<sup>7</sup>: C12N 9/00 (74) Agent: DÖRPER, Thomas; BASF Aktiengesellschaft, 67056 Ludwigshafen (DE).
- (21) International Application Number: PCT/EP02/03418 (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LI, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (22) International Filing Date: 26 March 2002 (26.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
0107510.0 26 March 2001 (26.03.2001) GB (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (*for all designated States except US*): UNIVERSITY OF BRISTOL [DE/DE]; Long Ashton, Bristol BS41 9AF (GB).
- (72) Inventors; and  
(75) Inventors/Applicants (*for US only*): NAPIER, Johnathan, A. [GB/GB]; 40 The Nursery, Bedminster, Bristol BS3 3EB (GB). LAZARUS, Colin, M. [GB/GB]; 119 York Road, Montpelier, Bristol BS6 5QG (GB). QI, Baoxiu [GB/GB]; 4 Cumberland House, Norfolk Crescent, Bath BA1 2BG (GB). LERCHL, Jens [DE/SE]; Onsjövägen 17, S-26831 Svalöv (SE).
- Published:  
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

WO 02/077213 A2

(54) Title: NEW ELONGASE GENE AND PRODUCTION OF Δ<sup>9</sup>-POLYUNSATURATED FATTY ACIDS

(57) Abstract: The invention relates to a new elongase gene having the sequence SEQ ID NO: 1 or its derivatives, to a gene construct comprising this sequence or its derivatives and to its use. The inventive nucleic acid sequence encodes a polypeptide which elongates α-linolenic acid (C<sub>18</sub>:3 d9, 12, 15?) by at least two carbon atoms whereas γ-linolenic acid (C<sub>18</sub>:3 d6, 9, 12?) is not elongated. The invention additionally relates to vectors or organisms comprising an elongase gene having the sequence SEQ ID NO: 1 or its derivatives. The invention further relates to a process for the production of polyunsaturated fatty acids (=PUFAs) with an organism which comprises the elongase gene and which organism produces high amounts of oils and especially oils with a high content of unsaturated fatty acids. Additionally the invention relates to an oil and/or fatty acid composition with a higher content of polyunsaturated C<sub>20</sub> or C<sub>22</sub> fatty acids with at least two double bonds and/or to a triacylglycerol composition with a higher content of said polyunsaturated fatty acids.

New elongase gene and production of  $\Delta 9$ -polyunsaturated fatty acids

## 5 Field of the Invention

This invention relates to a new elongase gene having the sequence SEQ ID NO: 1 or its derivatives, to a gene construct comprising this sequence or its derivatives and to its use. The inventive  
10 nucleic acid sequence encodes a polypeptide which elongates  $\alpha$ -linolenic acid ( $C_{18:3} \Delta 9, 12, 15$ ) by at least two carbon atoms whereas  $\gamma$ -linolenic acid ( $C_{18:3} \Delta 6, 9, 12$ ) is not elongated. The invention additionally relates to vectors or organisms comprising an elongase gene having the sequence SEQ ID NO: 1 or its  
15 derivatives.

The invention further relates to a process for the production of polyunsaturated fatty acids (= PUFAs) with an organism which comprises the elongase gene and which organism produces high  
20 amounts of oils and especially oils with a high content of unsaturated fatty acids. Additionally the invention relates to an oil and/or fatty acid composition with a higher content of polyunsaturated  $C_{20}$ - or  $C_{22}$ -fatty acids with at least two double bonds and/or to a triacylglycerol composition with a higher  
25 content of said polyunsaturated fatty acids.

## Background of the Invention

Certain products and by-products of naturally-occurring metabolic  
30 processes in cells have utility in a wide array of industries, including the food, feed, cosmetics, and pharmaceutical industries. These molecules, collectively termed 'fine chemicals', also include lipids and fatty acids whereof one representative class of molecules is polyunsaturated fatty acids.  
35 Polyunsaturated fatty acids (= PUFAs) are added for example to infant formulas to create a higher nutrition value of such formulas. PUFAs have for example a positive influence on the cholesterol level of the blood in humans and therefore are useful in the protection against heart diseases. Fine chemicals  
40 and polyunsaturated fatty acids (= PUFAs) can be isolated from animal sources such as for example fish or produced with microorganisms through the large-scale fermentation of microorganisms developed to produce and accumulate or secrete large quantities of one or more desired molecules.

45

## 2

Particularly useful microorganisms for the production of PUFAs are microorganisms such as the algae *Isochrysis galbana*, *Phaedactylum tricornutum* or *Cryptothecodinium* species, ciliates like *Stylonychia* or *Colpidium*, fungi like *Mortierella*,  
5 *Entomophthora*, *Mucor* or *Thrausto-/Schizochytrium* species. Through strain selection, a number of mutant strains of the respective microorganisms have been developed which produce an array of desirable compounds including PUFAs. However, selection of strains improved for the production of a particular molecule  
10 is a time-consuming and difficult process.

Alternatively the production of fine chemicals can be most conveniently performed via the large scale production of plants developed to produce aforementioned PUFAs. Particularly well  
15 suited plants for this purpose are oilseed plants containing high amounts of lipid compounds such as rapeseed, canola, linseed, soybean, sunflower, borage and evening primrose. But also other crop plants containing oils or lipids and fatty acids are well suited as mentioned in the detailed description of this  
20 invention. Through conventional breeding, a number of mutant plants have been developed which produce an array of desirable lipids and fatty acids, cofactors and enzymes. However, selection of new plant cultivars bred for the production of a particular molecule is a time-consuming and difficult process or even  
25 impossible if the compound does not naturally occur in the respective oil crop as in the case of C<sub>20</sub> and higher carbon chain polyunsaturated fatty acids.

### Summary of the Invention

30 This invention provides a novel nucleic acid molecule as described in SEQ ID NO: 1 which may be used to modify oils, fatty acids, lipids, lipid derived compounds and most preferred to produce polyunsaturated fatty acids.

35 Microorganisms such as *Mortierella*, *Entomophthora*, *Mucor*, *Cryptothecodinium* as well as other algae and fungi and plants, especially oilseed plants, are commonly used in industry for the large-scale production of a variety of fine chemicals.

40 Given the availability of cloning vectors and techniques for genetic manipulation of the abovementioned microorganisms and ciliates such as disclosed in WO 98/01572 or algae and related organisms such as *Phaeodactylum tricornutum* described in  
45 Falciatore et al., 1999, *Marine Biotechnology* 1 (3):239-251 as well as Dunahay et al. 1995, *Genetic transformation of diatoms*, *J. Phycol.* 31:10004-1012 and references therein, the nucleic

acid molecules of the invention may be utilized in the genetic engineering of these organisms to make them better or more efficient producers of one or more fine chemicals. This improved production or efficiency of production of a fine chemical may  
5 be due to a direct effect of manipulation of a gene of the invention, or it may be due to an indirect effect of such manipulation.

Mosses and algae are the only known plant systems that produce  
10 considerable amounts of polyunsaturated fatty acids like arachidonic acid (ARA) and/or eicosapentaenoic acid (EPA) and/or docosahexaenoic acid (DHA). Therefore nucleic acid molecules originating from an alga like *Isochrysis galbana* are especially suited to modify the lipid and PUFA production system in a  
15 host, especially in microorganisms such as the abovementioned microorganisms and plants such as oilseed plants, for example rapeseed, canola, linseed, soybean, sunflower, borage. Furthermore nucleic acids from the alga *Isochrysis galbana* can be used to identify those DNA sequences and enzymes in other species  
20 which are useful to modify the biosynthesis of precursor molecules of PUFAs in the respective organisms.

The alga *Isochrysis galbana* shares a high degree of homology on the DNA sequence and polypeptide levels with other algae allowing  
25 the use of heterologous screening of DNA molecules with probes evolving from other algae or organisms, thus enabling the derivation of a consensus sequence suitable for heterologous screening or functional annotation and prediction of gene functions in third species. The ability to identify such  
30 functions can therefore have significant relevance, e.g., prediction of substrate specificity of enzymes. Further, these nucleic acid molecules may serve as reference sequences for the mapping of other algae or for the derivation of PCR primers.

35 These novel nucleic acid molecules can encode proteins referred to herein as PUFA specific elongases (PSEs or singular PSE). These PSEs are capable of, for example, performing a function involved in the metabolism (e.g., the biosynthesis or degradation) of compounds necessary for lipid or fatty acid  
40 biosynthesis like PUFAs, or of assisting in the transmembrane transport of one or more lipid/fatty acid compounds either into or out of the cell.

In the present application we show the function of one of these  
45 sequences in more detail. We have isolated for the first time a functionally active gene encoding a highly specific elongase activity suitable to produce long chain unsaturated fatty acids



from  $\alpha$ -linolenic acid ( $C_{18:3} \Delta^9, 12, 15$ ) while  $\gamma$ -linolenic acid ( $C_{18:3} \Delta^6, 9, 12$ ) is not elongated. We will herein therefore refer to an ASE ("alpha-linolenic acid specific elongase") gene or protein thus representing an enzymatic activity leading to the  
5 elongation of omega-3 fatty acids or delta-9 desaturated long chain polyunsaturated fatty acids at least by two carbon atoms. Other publications and patents have not been able before to show a functionally active ASE gene that is specific for  $\alpha$ -linolenic acid (ALA) and which does not accept  $\gamma$ -linolenic acid (GLA) as  
10 a substrate though there are various patent applications known showing the elongation of short or medium chain saturated fatty acids (WO 98/46776 and US 5,475,099). WO 00/12720 describes various PSEs from various organisms but none of the described genes was shown to be specific for ALA while discriminating  
15 against GLA. Genes shown to encode PSEs from WO 00/12720 all accept GLA as a substrate, hence these enzymes lead to the elongation of  $\Delta^6$  desaturated long chain polyunsaturated fatty acids but not of  $\Delta^9$  desaturated fatty acids as disclosed in the present invention.

20 The unique feature of the ASE disclosed in the present invention is important as resulting products are limited to desired products while being free of unwanted PUFA molecules such as those resulting from the elongation of GLA.

25 WO 99/64616, WO 98/46763, WO 98/46764, WO 98/46765 describe the production of PUFAs in transgenic plants showing the cloning and functional expression of respective  $\Delta^{12}$ -,  $\Delta^5$ - or  $\Delta^6$ -desaturase activities from several sources lacking the demonstration of an  
30 ASE encoding gene or an  $\alpha$ -linolenic acid specific  $\Delta^6$ -desaturase gene and functional activity necessary for the production of eicosapentaenoic acid and related precursors from ALA.

For the production of PUFAs it is necessary that the  
35 polyunsaturated fatty acid molecules such as the preferred  $C_{18}$  fatty acids are elongated by at least two carbon atoms by the enzymatic activity of an elongase. The nucleic acid sequence of the invention encodes the first elongase derived from a plant which has the ability to elongate  $C_{18:3} \Delta^6, 9, 12$  fatty acids with  
40 at least two double bonds, preferably three double bonds, in the fatty acid by at least two carbon atoms. After one round of elongation this enzymatic activity leads to  $C_{20}$  fatty acids and after two, three and four rounds of elongation to  $C_{22}$ ,  $C_{24}$  or  $C_{26}$  fatty acids. With the elongase of the invention it is also  
45 possible to synthesize longer PUFAs. The activity of the elongase of the invention leads preferably to  $C_{20}$  and/or  $C_{22}$  fatty acids with at least two double bonds in the fatty acid molecule,

preferably with three or four double bonds, particularly preferably three double bonds, in the fatty acid molecule. Preferred fatty acid molecules of the elongation are fatty acid molecules with a double bond in  $\Delta 9$ -position. After the elongation with the inventive enzyme has taken place further desaturation steps might occur. Therefore the products of the elongase activity and the possible further desaturation lead to preferred PUFAs with a higher desaturation grade such as docosadienoic acid, arachidonic acid,  $\omega 6$ -eicosatrienoic acid, dihomo- $\gamma$ -linolenic acid, eicosapentenoic acid,  $\omega 3$ -eicosatrienoic acid,  $\omega 3$ -eicosatetraenoic acid, docosapentaenoic acid or docosahexaenoic acid. A particularly preferred fatty acid is the elongation product stearidonic acid of the  $\alpha$ -linolenic acid ( $C_{18:3} \Delta 9, 12, 15$ ). Substrates of the enzymatic activity of the invention are preferably  $\Delta 9$  desaturated fatty acids which have the first double bond in  $\Delta 9$ -position such as axillarenic acid, vernolic acid ( $C_{18}, \Delta 9_{cis}, 12-13_{epoxy}$ ), conjugated linoleic acid ( $C_{18}, \Delta 9_{cis}, 11_{trans}$ ), steric acid ( $C_{18}, \Delta 9$ -acetylenic),  $\alpha$ -parinaric acid ( $C_{18}, \Delta 9_{cis}, 11_{trans}, 13_{trans}$ ), palmitoleic acid ( $C_{18}, \Delta 9_{cis}$ ), linoleic acid or  $\alpha$ -linolenic acid. Preferred substrates are linoleic acid and/or  $\alpha$ -linolenic acid. The fatty acids with at least two double bonds in the fatty acid can be elongated by the enzymatic activity of the invention in the form of the free fatty acids, the acyl-CoA-fatty acids, alkyl esters of the fatty acids or in the form of the esters such as phospholipids, glycolipids, sphingolipids, phosphoglycerides, monoacylglycerides, diacylglycerides or triacylglycerides, preferably in the form of free fatty acid or the acyl-CoA-fatty acids.

Given the availability of cloning vectors for use in plants and plant transformation, such as those published in and cited therein: Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), chapter 6/7, p. 71-119 (1993); F.F. White, Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds.: Kung and R. Wu, Academic Press, 1993, 15-38; B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds.: Kung and R. Wu, Academic Press (1993), 128-143; Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205-225), the nucleic acid molecules of the invention may be utilized in the genetic engineering of a wide variety of plants to make them better or more efficient producers of one or more lipid derived products such as PUFAs. This improved production or efficiency of production of a lipid derived product such as PUFAs may be due to a direct effect of manipulation of a

gene of the invention, or it may be due to an indirect effect of such manipulation.

There are a number of mechanisms by which the alteration of  
5 an ASE protein of the invention may directly affect the yield, production, and/or efficiency of production of a fine chemical from an oilseed plant or microorganism due to such an altered protein. The ASE protein or gene may be increased in number or activity such that greater quantities of these compounds are  
10 produced or the compound may be produced de novo as the organisms were lacking this activity and capacity of biosynthesis before the introduction of the respective gene.

The introduction of an ASE gene into an organism or cell may not  
15 just increase the biosynthetic flux into an end product, it may also increase or create de novo the respective triacylglycerol composition. Similarly, other genes involved in the import of nutrients necessary for the biosynthesis of one or more fine chemicals (e.g., fatty acids, polar and neutral lipids) may be  
20 increased in number or activity such that these precursors, cofactors, or intermediate compounds are increased in concentration within the cell or within the storing compartment, thus increasing further the capability of the cell to produce PUFAs as described below. Fatty acids and lipids themselves are  
25 desirable fine chemicals; by optimizing the activity or increasing the number of one or more ASEs which participate in the biosynthesis of these compounds, or by impairing the activity of one or more ASEs which are involved in the degradation of these compounds, it may be possible to increase the yield,  
30 production, and/or efficiency of production of fatty acid and lipid molecules from plants or microorganisms.

The mutagenesis of the ASE gene of the invention may also result in an ASE protein having altered activities which directly or  
35 indirectly impact the production of one or more desired fine chemicals. For example the ASE gene of the invention may be increased in number or activity such that the normal metabolic wastes or byproducts of the cell (possibly increased in quantity due to the overproduction of the desired fine chemical) are  
40 efficiently exported before they are able to destroy other molecules or processes within the cell (which would decrease the viability of the cell) or to interfere with fine chemical biosynthetic pathways (which would decrease the yield, production, or efficiency of production of the desired fine  
45 chemical). Further, the relatively large intracellular quantities of the desired fine chemical may in themselves be toxic to the cell or may interfere with enzyme feedback mechanisms such as

allosteric regulation; for example, by increasing the activity or number of other downstream enzymes or detoxifying enzymes of the PUFA pathway, it might increase the allocation of the PUFA into the triacylglycerol fraction, one might increase the viability of seed cells, in turn leading to a better development of cells in the culture or in a seed producing the desired fine chemical. The ASE gene of the invention may also be manipulated such that the relative amounts of different lipid and fatty acid molecules are produced. This may have a profound effect on the lipid composition of the membrane of the cell and would create novel oils in addition to the occurrence of newly synthesized PUFAs. Since each type of lipid has different physical properties, an alteration in the lipid composition of a membrane may significantly alter membrane fluidity. Changes in membrane fluidity can impact the transport of molecules across the membrane, as well as the integrity of the cell, both of which have a profound effect on the production of fine chemicals. In plants these changes can moreover also influence other characteristics like tolerance towards abiotic and biotic stress conditions.

Biotic and abiotic stress tolerance is a general trait desired to be inherited into a wide variety of plants like maize, wheat, rye, oats, triticale, rice, barley, soybean, peanut, cotton, rapeseed and canola, manihot, pepper, sunflower and tagetes, solanaceous plants like potato, tobacco, eggplant, and tomato, Vicia species, pea, alfalfa, bushy plants (coffee, cacao, tea), Salix species, trees (oil palm, coconut) and perennial grasses and forage crops. These crop plants are also preferred target plants for genetic engineering as one further embodiment of the present invention. Particularly preferred plants of the invention are oilseed plants such as soybean, peanut, rapeseed, canola, sunflower, safflower, trees (oil palm, coconut) or crops such as maize, wheat, rye, oats, triticale, rice, barley, alfalfa or bushy plants (coffee, cacao, tea).

In another embodiment, the isolated nucleic acid molecule encodes a protein or portion thereof wherein the protein or portion thereof includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of sequence SEQ ID NO: 2 such that the protein or portion thereof maintains an ASE activity. Preferably, the protein or portion thereof encoded by the nucleic acid molecule maintains the ability to participate in the metabolism of compounds necessary for the construction of PUFAs or cellular membranes of plants or in the transport of molecules across these membranes. In one embodiment, the protein encoded by the nucleic acid molecule is at least about 50%,



- preferably at least about 60%, and more preferably at least about 70%, 80%, or 90% and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an amino acid sequence of sequence SEQ ID NO: 2. In another preferred embodiment, the protein is a full length *Isochrysis galbana* protein which is substantially homologous to an entire amino acid sequence of SEQ ID NO: 2 (derived from an open reading frame shown in SEQ ID NO: 1).
- 10 Accordingly, one aspect of the invention pertains to isolated nucleic acid molecules (e.g., cDNAs) comprising a nucleotide sequence encoding an ASE protein or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection or amplification of
- 15 ASE-encoding nucleic acids (e.g., DNA or mRNA). In particularly preferred embodiments, the nucleic acid molecule comprises one of the nucleotide sequences set forth in SEQ ID NO: 1, derivatives of said sequence or the coding region or a complement or enzymatically active part thereof. In other particularly
- 20 preferred embodiments, the isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes to or is at least about 50%, preferably at least about 60%, more preferably at least about 70%, 80% or 90%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more
- 25 homologous to a nucleotide sequence as in SEQ ID NO: 1, derivatives or a portion thereof. In other preferred embodiments, the isolated nucleic acid molecule encodes an amino acid sequence as set forth in SEQ ID NO: 2. The preferred ASE gene of the present invention also preferably possesses at least one of the
- 30 ASE activities described herein.

In another preferred embodiment, the isolated nucleic acid molecule is derived from *Isochrysis galbana* and encodes a protein (e.g., an ASE fusion protein) which includes a biologically

35 active domain which is at least about 50% or more homologous to one of the amino acid sequences of SEQ ID NO: 2 and is able to participate in the metabolism of compounds necessary for the construction of cellular membranes or in the transport of molecules across these membranes, or has one or more of the

40 activities set forth in Tab. 2, and which also includes heterologous nucleic acid sequences encoding a heterologous polypeptide or regulatory regions.

In another embodiment, the isolated nucleic acid molecule is at

45 least 15 nucleotides in length and hybridizes under stringent conditions with a nucleic acid molecule comprising a nucleotide sequence of SEQ ID NO: 1. Preferably, the isolated nucleic

acid molecule corresponds to a naturally-occurring nucleic acid molecule. More preferably, the isolated nucleic acid encodes a naturally-occurring *Isochrysis galbana* ASE, or a biologically active portion thereof.

5

Another aspect of the invention pertains to vectors, e.g., recombinant expression vectors, containing the nucleic acid molecules of the invention, and host cells into which such vectors have been introduced, especially microorganisms, plant  
10 cells, plant tissue, plant organs or whole plants. In one embodiment, such a host cell is a cell capable of storing fine chemical compounds, especially PUFAs, in order to isolate the desired compound from harvested material. The compound (oils, lipids, triacyl glycerides, fatty acids) or the ASE can then be  
15 isolated from the medium or the host cell, which in plants are cells containing and storing fine chemical compounds, most preferably cells of storage tissues like seed coats, tubers, epidermal and seed cells.

20 Yet another aspect of the invention pertains to an isolated ASE gene shown in SEQ ID NO: 1 or a portion thereof, e.g., a biologically active portion thereof. In a preferred embodiment, the isolated ASE or portion thereof can participate in the metabolism of compounds necessary for the construction of  
25 cellular membranes in a microorganism or a plant cell, or in the transport of molecules across its membranes. In another preferred embodiment, the isolated ASE or portion thereof is sufficiently homologous to an amino acid sequence of SEQ ID NO: 2 such that the protein or portion thereof maintains the ability to  
30 participate in the metabolism of compounds necessary for the construction of cellular membranes in microorganisms or plant cells, or in the transport of molecules across these membranes.

Hence in another preferred embodiment, the alga *Isochrysis*  
35 *galbana* can be used to show the function of a moss gene using homologous recombination based on the nucleic acids described in this invention.

Still another aspect of the invention pertains to an isolated ASE  
40 gene or a portion, e.g., a biologically active portion, thereof. In a preferred embodiment, the isolated ASE or portion thereof can participate in the metabolism of compounds necessary for the construction of cellular membranes in a microorganism or a plant cell, or in the transport of molecules across its membranes.  
45 In another preferred embodiment, the isolated ASE or portion thereof is sufficiently homologous to an amino acid sequence of SEQ ID NO: 2 such that the protein or portion thereof maintains

the ability to participate in the metabolism of compounds necessary for the construction of cellular membranes in microorganisms or plant cells, or in the transport of molecules across these membranes.

5

The invention also provides an isolated preparation of an ASE. In preferred embodiments, the ASE gene comprises an amino acid sequence of SEQ ID NO: 2. In another preferred embodiment, the invention pertains to an isolated full length protein which is  
10 substantially homologous to an entire amino acid sequence of SEQ ID NO: 2 (encoded by an open reading frame set forth in SEQ ID NO: 1). In another embodiment, the protein is at least about 50%, preferably at least about 60%, and more preferably at least about 70%, 80%, or 90%, and most preferably at least about 95%, 96%,  
15 97%, 98%, or 99% or more homologous to an entire amino acid sequence of SEQ ID NO: 2. In other embodiments, the isolated ASE comprises an amino acid sequence which is at least about 50% or more homologous to one of the amino acid sequences of SEQ ID NO: 2 and is able to participate in the metabolism of compounds  
20 necessary for the construction of fatty acids in a microorganism or a plant cell, or in the transport of molecules across these membranes, or has one or more of the PUFA elongating activities, thus meaning the elongation of C18 carbon chains being desaturated bearing at least two double bond positions.

25

Alternatively, the isolated ASE can comprise an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, or is at least about 50%, preferably at least about 60%, more  
30 preferably at least about 70%, 80%, or 90%, and even more preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous, to a nucleotide sequence of SEQUENCE ID NO: 1. It is also preferred that the preferred forms of ASEs also have one or more of the ASE activities described herein.

35

The ASE polypeptide, or a biologically active portion thereof, can be operatively linked to a non-ASE polypeptide to form a fusion protein. In preferred embodiments, this fusion protein has an activity which differs from that of the ASE alone. In other  
40 preferred embodiments, this fusion protein participates in the metabolism of compounds necessary for the synthesis of lipids and fatty acids, cofactors and enzymes in microorganisms or plants, or in the transport of molecules across the membranes of plants. In particularly preferred embodiments, integration of this fusion  
45 protein into a host cell modulates production of a desired compound from the cell. In a preferred embodiment such fusion proteins also contain  $\Delta 4$ -,  $\Delta 5$ - or  $\Delta 8$ -desaturase activities alone

or in combination. Especially a  $\Delta 8$ -desaturase from *Euglena gracilis* described in WO 00/34439 and a  $\Delta 5$ -desaturase gene described in US 6,051,754 (*M. alpina*), GB9814034.6 (*C. elegans*) or a  $\Delta 12$ - and  $\Delta 15$ -desaturase gene described in US 5,850,026 are  
5 suitable genes for coexpression with an ASE gene of the present invention. None of the cited patents shows coexpression with an ASE gene described in the present invention.

Another aspect of the invention pertains to a method for  
10 producing a fine chemical. This method involves either the culturing of a suitable microorganism or culturing plant cells, tissues, organs or whole plants containing a vector directing the expression of an ASE nucleic acid molecule of the invention, such that a fine chemical is produced. In a preferred embodiment, this  
15 method further includes the step of obtaining a cell containing such a vector, in which a cell is transformed with a vector directing the expression of an ASE nucleic acid. In another preferred embodiment, this method includes the step of recovering the fine chemical from the culture. In a particularly preferred  
20 embodiment, the cell is from an alga such as *Phaeodactylum*, ciliates such as *Colpidium* or *Stylonichia*, fungi such as *Mortierella* or *Thraustochytrium* or *Schizochytrium* or from oilseed plants as mentioned above.

25 Another aspect of the invention pertains to methods for modulating production of a molecule from a microorganism. Such methods include contacting the cell with an agent which modulates ASE activity or ASE nucleic acid expression such that a cell associated activity is altered relative to this same activity in  
30 the absence of the agent. In a preferred embodiment, the cell is modulated for one or more metabolic pathways for lipids and fatty acids, cofactors and enzymes or is modulated for the transport of compounds across such membranes, such that the yields or rates of production of a desired fine chemical by this microorganism are  
35 improved. The agent which modulates ASE activity can be an agent which stimulates ASE activity or ASE nucleic acid expression. Examples of agents which stimulate ASE activity or ASE nucleic acid expression include small molecules, active ASEs, and nucleic acids encoding ASEs that have been introduced into the cell.  
40 Examples of agents which inhibit ASE activity or expression include small molecules and antisense ASE nucleic acid molecules.

Another aspect of the invention pertains to methods for modulating yields of a desired compound from a cell, involving  
45 the introduction of a wild-type or mutant ASE gene into a cell, the gene being either maintained on a separate plasmid or integrated into the genome of the host cell. If integrated into



the genome, such integration can be random, or it can take place by recombination such that the native gene is replaced by the introduced copy, causing the production of the desired compound from the cell to be modulated or by using a gene in trans such  
5 that the gene is functionally linked to a functional expression unit containing at least a sequence facilitating the expression of a gene and a sequence facilitating the polyadenylation of a functionally transcribed gene.

- 10 In a preferred embodiment, said yields are modified. In another preferred embodiment, said desired chemical is increased while unwanted disturbing compounds can be decreased. In a particularly preferred embodiment, said desired fine chemical is a lipid or fatty acid, cofactor or enzyme. In an especially preferred  
15 embodiment, said chemical is a polyunsaturated fatty acid. More preferably it is chosen from arachidonic acid (ARA), eicosapentaenoic acid (EPA) or docosahexaenoic acid (DHA).

#### Detailed Description of the Invention

20

- The present invention provides ASE nucleic acid and protein molecules which are involved in the metabolism of lipids and fatty acids, PUFA cofactors and enzymes in the alga *Isochrysis galbana* or in the transport of lipophilic compounds across  
25 membranes. The molecules of the invention may be utilized in the modulation of production of fine chemicals from microorganisms, such as ciliates such as *Colpidium* or *Stylonichia*, fungi such as *Mortierella* or *Thraustochytrium* or *Schizochytrium*, algae such as *Phaeodactylum*, and/or plants like maize, wheat, rye, oats,  
30 triticale, rice, barley, soybean, peanut, cotton, Brassica species like rapeseed, canola and turnip rape, linseed, pepper, sunflower, borage, evening primrose and tagetes, solanaceous plants like potato, tobacco, eggplant, and tomato, Vicia species, pea, manihot, alfalfa, bushy plants (coffee, cacao, tea), Salix  
35 species, trees (oil palm, coconut) and perennial grasses and forage crops either directly (e.g., where overexpression or optimization of a fatty acid biosynthesis protein has a direct impact on the yield, production, and/or efficiency of production of the fatty acid from modified organisms), or may have an  
40 indirect impact which nonetheless results in an increase of yield, production, and/or efficiency of production of a desired compound or decrease of undesired compounds (e.g., where modulation of the metabolism of lipids and fatty acids, cofactors and enzymes results in alterations in the yield, production,  
45 and/or efficiency of production or the composition of desired compounds within the cells, which in turn may impact the

production of one or more fine chemicals). Aspects of the invention are further explicated below.

#### I. Fine Chemicals and PUFAs

5

The term 'fine chemical' is art-recognized and includes molecules produced by an organism which have applications in various industries, such as, but not limited to, the pharmaceutical, agriculture, food, feed and cosmetics industries. Such compounds  
10 also include lipids, fatty acids, cofactors and enzymes etc. (as described e.g. in Kuninaka, A. (1996) Nucleotides and related compounds, p. 561-612, in Biotechnology vol. 6, Rehm et al., eds. VCH: Weinheim, and references contained therein), lipids, both saturated and polyunsaturated fatty acids (e.g., arachidonic  
15 acid), vitamins and cofactors (as described in Ullmann's Encyclopedia of Industrial Chemistry, vol. A27, Vitamins, p. 443-613 (1996) VCH Weinheim and references therein; and Ong, A.S., Niki, E. & Packer, L. (1995) Nutrition, Lipids, Health, and Disease Proceedings of the UNESCO/Confederation of Scientific and  
20 Technological Associations in Malaysia, and the Society for Free Radical Research, Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press, (1995)), enzymes, and all other chemicals described by Gutcho (1983) in Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 and references therein. The  
25 metabolism and uses of certain of these fine chemicals are further explicated below.

The combination of various precursor molecules and biosynthetic enzymes results in the production of different fatty acid  
30 molecules, which has a profound effect on the composition of the membrane. It can be assumed that PUFAs will not just be incorporated into triacylglycerol but also into membrane lipids.

The synthesis of membranes is a well-characterized process  
35 involving a number of components including lipids as part of the bilayer membrane. The production of new fatty acids such as PUFAs may therefore create new characteristics of membrane functions within a cell or organism.

40 Cellular membranes serve a variety of functions in a cell. First and foremost, a membrane differentiates the contents of a cell from the surrounding environment, thus giving integrity to the cell. Membranes may also serve as barriers to the influx of hazardous or unwanted compounds, and also to the efflux of  
45 desired compounds.

For more detailed descriptions and implications of membranes and involved mechanisms see: Bamberg, E. et al., (1993) Charge transport of ion pumps on lipid bilayer membranes, Q. Rev. Biophys. 26: 1-25; Gennis, R.B. (1989) Pores, Channels and  
5 Transporters, in: Biomembranes, Molecular Structure and Function, Springer: Heidelberg, p. 270-322; and Nikaido, H. and Saier, H. (1992) Transport proteins in bacteria: common themes in their design, Science 258: 936-942, and references contained within each of these references.

10

Lipid synthesis may be divided into two parts: the synthesis of fatty acids and their attachment to sn-glycerol-3-phosphate, and the addition or modification of a polar head group. Typical lipids utilized in membranes include phospholipids, glycolipids,  
15 sphingolipids, and phosphoglycerides. Fatty acid synthesis begins with the conversion of acetyl CoA either to malonyl CoA by acetyl CoA carboxylase, or to acetyl-ACP by acetyltransacylase. Following a condensation reaction, these two product molecules together form further intermediates, which are converted by  
20 a series of condensation, reduction and dehydration reactions to yield a saturated fatty acid molecule having the desired chain length. The production of unsaturated fatty acids from such molecules is catalyzed by specific desaturases either aerobically, with the help of molecular oxygen, or anaerobically  
25 (for reference on fatty acid synthesis in microorganisms, see F.C. Neidhardt et al. (1996) E. coli and Salmonella. ASM Press: Washington, D.C., p. 612-636 and references contained therein; Lengeler et al. (eds) (1999) Biology of Prokaryotes. Thieme: Stuttgart, New York, and references contained therein; and  
30 Magnuson, K. et al., (1993) Microbiological Reviews 57: 522-542, and references contained therein).

Preferred precursors for the inventive PUFA biosynthesis process are linoleic and linolenic acid. These C<sub>18</sub> carbon fatty acids  
35 have to be elongated to C<sub>20</sub> and C<sub>22</sub> in order to obtain eicosa and docosa chain type fatty acids. With the aid of various desaturases such as enzymes with  $\Delta^6$ - or  $\Delta^8$ - and  $\Delta^5$ - and  $\Delta^4$ -desaturase activity eicosapentaenoic acid and docosahexaenoic acid as well as various other long chain PUFAs can be obtained,  
40 extracted and used for various purposes, for example in food and feed applications.

For the production of long chain PUFAs it is necessary as mentioned above that the polyunsaturated C<sub>18</sub> fatty acids are  
45 elongated by at least two carbon atoms by the enzymatic activity of the inventive elongase. The nucleic acid sequence of the invention encodes the first plant elongase which has the ability

## 15

to elongate  $\alpha$ -linolenic acid ( $C_{18:3} \Delta^9, 12, 15$ ) by at least two carbon atoms but not  $\gamma$ -linolenic acid ( $C_{18:3} \Delta^6, 9, 12$ ).

Furthermore fatty acids have to be transported and incorporated  
5 into the triacylglycerol storage lipid subsequent to various modifications. Another essential step in lipid synthesis is the transfer of fatty acids onto the polar head groups by, for example, glycerol-phosphate-acyltransferase (see Frentzen, 1998, Lipid, 100(4-5):161-166).

10

For publications on plant fatty acid biosynthesis, desaturation, lipid metabolism and membrane transport of lipoic compounds, beta-oxidation, fatty acid modification and cofactors, triacylglycerol storage and assembly including references therein  
15 see following articles: Kinney, 1997, Genetic Engineering, ed.: JK Setlow, 19:149-166; Ohlrogge and Browse, 1995, Plant Cell 7:957-970; Shanklin and Cahoon, 1998, Annu. Rev. Plant Physiol. Plant Mol. Biol., 49:611-641; Voelker, 1996, Genetic Engineering, ed.: JK Setlow, 18:111-13; Gerhardt, 1992, Prog. Lipid R.  
20 31:397-417; Gühnemann-Schäfer & Kindl, 1995, Biochim. Biophys Acta 1256:181-186; Kunau et al., 1995, Prog. Lipid Res. 34:267-342; Stymne et al 1993, in: Biochemistry and Molecular Biology of Membrane and Storage Lipids of Plants, Eds: Murata and Somerville, Rockville, American Society of Plant  
25 Physiologists, 150-158, Murphy & Ross 1998, Plant Journal. 13(1):1-16.

Vitamins, cofactors, and nutraceuticals such as PUFAs comprise a group of molecules which the higher animals have lost the ability  
30 to synthesize and so must ingest or which the higher animals cannot sufficiently produce on their own and so must ingest additionally, although they are readily synthesized by other organisms such as bacteria. The biosynthesis of these molecules in organisms capable of producing them, such as bacteria, has  
35 been largely characterized (Ullmann's Encyclopedia of Industrial Chemistry, Vitamins vol. A27, p. 443-613, VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E. & Packer, L. (1995) Nutrition, Lipids, Health, and Disease  
40 Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press: Champaign, IL X, 374 S).

45 These said molecules are either bioactive substances themselves, or are precursors of biologically active substances which may serve as electron carriers or intermediates in a variety of



metabolic pathways. Aside from their nutritive value, these compounds also have significant industrial value as coloring agents, antioxidants, and catalysts or other processing aids. (For an overview of the structure, activity, and industrial applications of these compounds, see, for example, Ullmann's Encyclopedia of Industrial Chemistry, Vitamins vol. A27, p. 443-613, VCH: Weinheim, 1996.). Polyunsaturated fatty acids have various functions and health benefit effects such as in coronary heart disease, inflammatory mechanisms, infant nutrition etc.

For publications and references see including references cited therein: Simopoulos 1999, Am. J. Clin. Nutr., 70 (3 Suppl):560-569, Takahata et al., Biosc. Biotechnol. Biochem, 1998, 62 (11):2079-2085, Willich und Winther, 1995, Deutsche Medizinische Wochenschrift, 120 (7):229 ff.

15

## II. Elements and Methods of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as ASE nucleic acid and protein molecules, which have an effect on the production of cellular membranes in *Isochrysis galbana* and influence for example the movement of molecules across such membranes. In one embodiment, the ASE molecules participate in the metabolism of compounds necessary for the construction of cellular membranes in microorganisms and plants, or directly influence the transport of molecules across these membranes. In a preferred embodiment, the activity of the ASE molecules of the present invention to regulate membrane component production and membrane transport has an impact on the production of a desired fine chemical by this organism. In a particularly preferred embodiment, the ASE molecules of the invention are modulated in activity, such that the microorganisms' or plants' metabolic pathways which the ASEs of the invention regulate are modulated in yield, production, and/or efficiency of production and the transport of compounds through the membranes is altered in efficiency, which either directly or indirectly modulates the yield, production, and/or efficiency of production of a desired fine chemical by microorganisms and plants.

The language, ASE or ASE polypeptide, includes proteins which participate in the metabolism of compounds necessary for the construction of cellular membranes in microorganisms and plants, or in the transport of molecules across these membranes. Examples of ASEs are disclosed in SEQ ID NO: 1 or its derivatives. The terms ASE gene or ASE nucleic acid sequence include nucleic acid sequences encoding an ASE, which consist of a coding region and also corresponding untranslated 5' and 3' sequence regions.

Examples of ASE genes include those set forth in SEQ ID NO: 1 and their derivatives. The terms production or productivity are art-recognized and include the concentration of the fermentation product (for example, the desired fine chemical) formed within a 5 given time and a given fermentation volume (e.g., kg product per hour per liter). The term efficiency of production includes the time required for a particular level of production to be achieved (for example, how long it takes for the cell to attain a particular throughput of a fine chemical). The term yield 10 or product/carbon yield is art-recognized and includes the efficiency of the conversion of the carbon source into the product (i.e., fine chemical). This is generally written as, for example, kg product per kg carbon source. By increasing the yield or production of the compound, the quantity of recovered 15 molecules, or of useful recovered molecules, of that compound in a given amount of culture over a given amount of time is increased. The terms biosynthesis or a biosynthetic pathway are art-recognized and include the synthesis of a compound, preferably an organic compound, by a cell from intermediate 20 compounds in what may be a multistep and highly regulated process. The terms degradation or a degradation pathway are art-recognized and include the breakdown of a compound, preferably an organic compound, by a cell to degradation products (generally speaking, smaller or less complex molecules) in what 25 may be a multistep and highly regulated process. The language metabolism is art-recognized and includes the totality of the biochemical reactions that take place in an organism. The metabolism of a particular compound, then, (e.g., the metabolism of a fatty acid) comprises the overall biosynthetic, 30 modification, and degradation pathways in a cell related to this compound.

In another embodiment, the ASE molecules of the invention are capable of modulating the production of a desired molecule, such 35 as a fine chemical, in microorganisms or plants. There are a number of mechanisms by which the alteration of an ASE of the invention may directly affect the yield, production, and/or efficiency of production of a fine chemical from a microorganism or plant strain incorporating such an altered protein. Those ASEs 40 involved in the transport of fine chemical molecules within or from the cell may be increased in number or activity such that greater quantities of these compounds are transported across membranes, from which they are more readily recovered and interconverted. Further, fatty acids and lipids themselves 45 are desirable fine chemicals; by optimizing the activity or increasing the number of one or more ASEs of the invention which participate in the biosynthesis of these compounds, or by

impairing the activity of one or more ASEs which are involved in the degradation of these compounds, it may be possible to increase the yield, production, and/or efficiency of production of fatty acid and lipid molecules from microorganisms or plants.

5

The mutagenesis of the ASE gene of the invention may also result in ASEs having altered activities which indirectly impact on the production of one or more desired fine chemicals from microorganisms or plants. For example, ASEs of the invention  
10 involved in the export of waste products may be increased in number or activity such that the normal metabolic wastes of the cell (possibly increased in quantity due to the overproduction of the desired fine chemical) are efficiently exported before they are able to damage molecules within the cell (which would  
15 decrease the viability of the cell) or to interfere with fine chemical biosynthetic pathways (which would decrease the yield, production, or efficiency of production of the desired fine chemical). Further, the relatively large intracellular quantities of the desired fine chemical may in themselves be toxic to the  
20 cell, so by increasing the activity or number of transporters able to export this compound from the cell, one may increase the viability of the cell in culture, in turn leading to a greater number of cells in the culture producing the desired fine chemical. The ASEs of the invention may also be manipulated such  
25 that the relative amounts of different lipid and fatty acid molecules are produced. This may have a profound effect on the lipid composition of the membrane of the cell. Since each type of lipid has different physical properties, an alteration in the lipid composition of a membrane may significantly alter membrane  
30 fluidity. Changes in membrane fluidity can impact the transport of molecules across the membrane, as well as the integrity of the cell, both of which have a profound effect on the production of fine chemicals from microorganisms and plants in large-scale fermentative culture. Plant membranes confer specific  
35 characteristics such as tolerance towards heat, cold, salt, drought and tolerance towards pathogens like bacteria and fungi. Modulating membrane compounds therefore can have a profound effect on the plants' fitness to survive under aforementioned stress parameters. This can happen either via changes in  
40 signaling cascades or directly via the changed membrane composition (for example see: Chapman, 1998, Trends in Plant Science, 3 (11):419-426) and influence signaling cascades (see Wang 1999, Plant Physiology, 120:645-651) or on cold tolerance as disclosed in WO 95/18222.

45

## 19

The isolated nucleic acid sequence of the invention is contained within the genome of an *Isochrysis galbana* strain as described in the Examples. The nucleotide sequence of the isolated *Isochrysis galbana* ASE cDNA and the predicted amino acid sequences of the *Isochrysis galbana* ASEs are shown in SEQ ID NO: 1 and 2, respectively.

A fragment of nucleic acid molecule in SEQ ID NO: 1 was isolated by polymerase chain reaction with the aid of degenerated oligonucleotides derived from other known elongase genes and a vector primer. A partial fragment was amplified further and used for the isolation of a full length cDNA containing sufficient sequence information representing a functionally active ASE gene. One clone contained a complete ASE gene showing weak homology to known elongase genes. The expression of the open reading frame in yeast unforeseeingly revealed an ASE gene specific activity. The enzyme elongates  $\Delta 9$ -fatty acids as shown in the Examples.

The present invention also pertains to proteins which have an amino acid sequence which is substantially homologous to an amino acid sequence of SEQ ID NO: 2. As used herein, a protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence is at least about 50% homologous to the selected amino acid sequence, e.g., the entire selected amino acid sequence. A protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence can also be least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, or 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to the selected amino acid sequence.

The ASE of the invention or a biologically active portion or fragment thereof can participate in the metabolism of compounds necessary for the construction of cellular membranes in microorganisms or plants, or in the transport of molecules across these membranes, or have one or more of the activities needed to elongate C18 PUFAs to yield in C<sub>22</sub> or C<sub>24</sub> PUFAs as well as related PUFAs.

Various aspects of the invention are described in further detail in the following subsections:

A. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode ASE polypeptides or biologically active



## 20

portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes or primers for the identification or amplification of ASE-encoding nucleic acid (e.g., ASE DNA). As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. This term also encompasses untranslated sequences located at both the 3' and 5' ends of the coding region of the gene: at least about 500, preferably at least about 400, more preferably at least about 350, 300, 250, 200, 150 and even more preferably at least about 100 nucleotides of sequence upstream from the 5' end of the coding region and at least about 1000, preferably at least about 500, more preferably at least about 400, 300, 250, 200, 150 and even more preferably 100, 80, 60, 40 or 20 nucleotides of sequence downstream from the 3' end of the coding region of the gene. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated ASE nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (e.g., an *Ischochrysis galbana* cell). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized; the language "substantially free of cellular material" includes preparations of nucleic acid molecules having less than about 30% (by dry weight) of other material such as proteins, polysaccharides etc. (also referred to herein as a "contaminating material"), more preferably less than about 20% of contaminating material, still more preferably less than about 10% of contaminating material, and most preferably less than about 5% of contaminating material.

One embodiment of the invention is an isolated nucleic acid derived from a plant encoding a polypeptide which elongates  $\alpha$ -linolenic acid ( $C_{18:3}$   $\Delta^9, 12, 15$ ) by at least two carbon atoms whereas  $\gamma$ -linolenic acid ( $C_{18:3}$   $\Delta^6, 9, 12$ ) is not elongated.

## 21

A further embodiment of the invention is an isolated nucleic acid comprising a nucleotide sequence which encodes a polypeptide which elongates  $\alpha$ -linolenic acid ( $C_{18:3}$   $\Delta^9, 12, 15$ ) by at least two carbon atoms whereas  $\gamma$ -linolenic acid ( $C_{18:3}$   $\Delta^6, 9, 12$ ) is not  
5 elongated, which nucleic acid is selected from the group consisting of

- a) a nucleic acid sequence depicted in SEQ ID NO: 1,
- 10 b) a nucleic acid sequence which encodes a polypeptide depicted in SEQ ID NO: 2,
- c) derivatives of the sequence depicted in SEQ ID NO: 1, which encodes polypeptides having at least 50% homology  
15 to the sequence encoding amino acid sequences depicted in SEQ ID NO: 2 and which sequences function as an elongase.

The abovementioned isolated nucleic acid of the invention is derived from organisms such as ciliates, fungi, algae or  
20 dinoflagellates which are able to synthesize PUFAs, preferably from plants, particularly preferably from the genus *Isochrysis* and most particularly preferably from *Isochrysis galbana*.

One aspect of the invention pertains to isolated nucleic acid  
25 molecules that encode ASE polypeptides or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes or primers for the identification or amplification of ASE-encoding nucleic acid (e.g., ASE DNA). As used herein, the term "nucleic acid molecule" is intended  
30 to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. This term also encompasses the untranslated sequence located at both the 3' and 5' ends of the coding region of the gene: at least about 100 nucleotides of  
35 sequence upstream from the 5' end of the coding region and at least about 20 nucleotides of sequence downstream from the 3' end of the coding region of the gene. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. An "isolated" nucleic acid molecule is one  
40 which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism  
45 from which the nucleic acid is derived. For example, in various embodiments, the isolated ASE nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb

of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (e.g., an *Isochrysis galbana* cell). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. Substantially free means that the [lacuna]

- 10 A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, an *Isochrysis galbana* ASE cDNA can be isolated from an
- 15 *Isochrysis galbana* library using all or a portion of SEQ ID NO: 1 as a hybridization probe and standard hybridization techniques (e.g., as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).
- 20 Moreover, a nucleic acid molecule encompassing all or a portion of one of the sequences of SEQ ID NO: 1 can be isolated by the polymerase chain reaction using oligonucleotide primers designed on the basis of this sequence or parts thereof, especially regions around his-box motifs, see Shanklin et al. (1994)
- 25 *Biochemistry* 33, 12787-12794 (e.g., a nucleic acid molecule encompassing all or a portion of one of the sequences of SEQ ID NO: 1 can be isolated by the polymerase chain reaction using oligonucleotide primers designed on the basis of this same sequence of SEQ ID NO: 1). For example, mRNA can be isolated from
- 30 plant cells (e.g., by the guanidinium-thiocyanate extraction procedure of Chirgwin et al. (1979) *Biochemistry* 18: 5294-5299) and cDNA can be prepared using reverse transcriptase (e.g., Moloney MLV reverse transcriptase, available from Gibco/BRL, Bethesda, MD; or AMV reverse transcriptase, available from
- 35 Seikagaku America, Inc., St. Petersburg, FL). Synthetic oligonucleotide primers for polymerase chain reaction amplification can be designed on the basis of one of the nucleotide sequence shown in SEQ ID NO: 1. A nucleic acid of the invention can be amplified using cDNA or, alternatively, genomic
- 40 DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to an ASE nucleotide sequence
- 45 can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

## 23

The cDNA shown in SEQUENCE ID NO: 1 comprises sequences encoding ASEs (i.e., the "coding region"), as well as 5' untranslated sequence and 3' untranslated sequence information. Alternatively, the nucleic acid molecule can comprise only the coding region of  
5 any of the sequences in SEQ ID NO: 1 or can contain whole genomic fragments isolated from genomic DNA.

The SEQUENCE ID NO: 2 is a translation of the coding region of the nucleotide sequence of nucleic acid molecule Ig\_ASE1 shown  
10 in SEQ ID NO: 1.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of one of the nucleotide sequences shown  
15 in SEQ ID NO: 1, or a portion thereof. A nucleic acid molecule which is complementary to one of the nucleotide sequences shown in SEQ ID NO: 1 is one which is sufficiently complementary to one of the nucleotide sequences shown in SEQ ID NO: 1 such that it can hybridize with one of the nucleotide sequences shown in  
20 SEQ ID NO: 1, thereby forming a stable duplex.

Homologs of the novel elongase nucleic acid sequence having the sequence SEQ ID NO: 1 mean, for example, allelic variants which have at least about 50-60%, preferably at least about 60-70%,  
25 more preferably at least about 70-80%, 80-90%, or 90-95%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homology to a nucleotide sequence shown in SEQ ID NO: 1 or its homologs, derivatives or analogs or portions thereof. In an additional preferred embodiment, an isolated nucleic acid  
30 molecule of the invention comprises a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, with one of the nucleotide sequences shown in SEQ ID NO: 1, or a portion thereof. Allelic variants comprise, in particular, functional variants which are obtainable by deletion, insertion or  
35 substitution of nucleotides from the sequence depicted in SEQ ID NO: 1, the intention being, however, that the enzymatic activity of the derived synthesized proteins advantageously be retained for the insertion of one or more genes. Proteins which have still the enzymatic activity of the elongase means proteins  
40 which have at least 10% of the original enzymatic activity, preferably 20%, particularly preferably 30%, most particularly preferably 40%, in comparison to the protein encoded by SEQ ID NO: 2.



## 24

Homologs of SEQ ID NO: 1 additionally mean, for example, bacterial, fungal or plant homologs, truncated sequences, single-stranded DNA or RNA of the coding and noncoding DNA sequence.

5

Homologs of SEQ ID NO: 1 also mean derivatives such as, for example, promoter variants. The promoters upstream of the indicated nucleotide sequences may be modified by one or more nucleotide exchanges, by insertion(s) and/or deletion(s) without, however, the functionality or activity of the promoters being impaired. It is additionally possible for the promoters to have their activity increased by modifying their sequence, or to be completely replaced by more active promoters even from heterologous organisms.

15

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the coding region of one of the sequences in SEQ ID NO: 1, for example a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of an ASE. The nucleotide sequences determined from the cloning of the ASE gene from *Isochrysis galbana* allow for the generation of probes and primers designed for use in identifying and/or cloning ASE homologs in other cell types and organisms, as well as ASE homologs from *Isochrysis galbana* or related species.

The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions with at least about 12, preferably about 16, more preferably about 25, 40, 50 or 75 consecutive nucleotides of a sense strand of one of the sequences set forth in SEQUENCE ID NO: 1, an antisense sequence of one of the sequences set forth in sequence ID NO: 1, or naturally occurring mutants thereof. Primers based on a nucleotide sequence of SEQUENCE ID NO: 1 can be used in PCR reactions to clone ASE homologs. Probes based on the ASE nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, e.g. the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a genomic marker test kit for identifying cells which misexpress an ASE, such as by measuring a level of an ASE-encoding nucleic acid in a sample of cells, e.g., detecting ASE mRNA levels or determining whether a genomic ASE gene has been mutated or deleted.

45

## 25

In one embodiment, the nucleic acid molecule of the invention encodes a protein or portion thereof which includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of SEQ ID NO: 2 such that the protein or portion thereof maintains the ability to participate in the metabolism of compounds necessary for the construction of cellular membranes in microorganisms or plants, or in the transport of molecules across these membranes. As used herein, the term "sufficiently homologous" refers to proteins or portions thereof which have amino acid sequences which include a minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain as an amino acid residue in one of the sequences of SEQ ID NO: 2) amino acid residues to an amino acid sequence of SEQ ID NO: 2 such that the protein or portion thereof is able to participate in the metabolism of compounds necessary for the construction of cellular membranes in microorganisms or plants, or in the transport of molecules across these membranes. Protein members of such membrane component metabolic pathways or membrane transport systems, as described herein, may play a role in the production and secretion of one or more fine chemicals. Examples of such activities are also described herein. Thus, the function of an ASE contributes either directly or indirectly to the yield, production, and/or efficiency of production of one or more fine chemicals. Examples of ASE substrate specificities of the catalytic activity are set forth in Tab. 2.

In another embodiment, derivatives of the nucleic acid molecule of the invention encode proteins which are at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to an entire amino acid sequence of SEQ ID NO: 2.

Portions of proteins encoded by the ASE nucleic acid molecules of the invention are preferably biologically active portions of one of the ASEs. As used herein, the term "biologically active portion of an ASE" is intended to include a portion, e.g., a domain/motif, of an ASE that participates in the metabolism of compounds necessary for the construction of cellular membranes in microorganisms or plants, or in the transport of molecules across these membranes, or has an activity as set forth in Tab. 2. To determine whether an ASE or a biologically active portion thereof can participate in the metabolism of compounds necessary for the construction of cellular membranes in microorganisms or plants, or in the transport of molecules across these membranes, an assay of enzymatic activity may be performed. Such assay methods are

well known to those skilled in the art, as detailed in Example 8 of the Examples.

Additional nucleic acid fragments encoding biologically active portions of an ASE can be prepared by isolating a portion of one of the sequences in SEQ ID NO: 2, expressing the encoded portion of the ASE or peptide (e.g., by recombinant expression in vitro) and assessing the activity of the encoded portion of the ASE or peptide.

10

The invention further encompasses nucleic acid molecules that differ from one of the nucleotide sequences shown in SEQ ID NO: 1 (and portions thereof) due to degeneracy of the genetic code and thus encode the same ASE as that encoded by the nucleotide sequences shown in SEQ ID NO: 1. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO: 2. In a further embodiment, the nucleic acid molecule of the invention encodes a full length *Isochrysis galbana* protein which is substantially homologous to an amino acid sequence of SEQ ID NO: 2 (encoded by an open reading frame shown in SEQ ID NO: 1).

In addition to the *Isochrysis galbana* ASE nucleotide sequences shown in SEQ ID NO: 1, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of ASEs may exist within a population (e.g., the *Isochrysis galbana* population). Such genetic polymorphisms in the ASE gene may exist among individuals within a population due to natural variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding an ASE, preferably an *Isochrysis galbana* ASE. Such natural variations can typically result in 1-5% variance in the nucleotide sequence of the ASE gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in ASE that are the result of natural variation and that do not alter the functional activity of ASEs are intended to be within the scope of the invention.

Nucleic acid molecules corresponding to natural variants and non-*Isochrysis galbana* homologs of the *Isochrysis galbana* ASE cDNA of the invention can be isolated based on their homology to *Isochrysis galbana* ASE nucleic acid disclosed herein using the *Isochrysis galbana* cDNA, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is

at least 15 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising a nucleotide sequence of SEQ ID NO: 1. In other embodiments, the nucleic acid is at least 25, 50, 100, 250 or more nucleotides in length. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are such that sequences at least about 65%, more preferably at least about 70%, and even more preferably at least about 75% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, nonlimiting example of stringent hybridization conditions is hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45 degree C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65 degree C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to a sequence of SEQ ID NO: 1 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein). In one embodiment, the nucleic acid encodes a natural Isochrysis galbana ASE.

In addition to naturally-occurring variants of the ASE sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into a nucleotide sequence of SEQ ID NO: 1, thereby leading to changes in the amino acid sequence of the encoded ASE, without altering the functional ability of the ASE. For example, nucleotide substitutions leading to amino acid substitutions at "nonessential" amino acid residues can be made in a sequence of SEQ ID NO: 1. A "nonessential" amino acid residue is a residue that can be altered from a wild-type sequence of one of the ASEs (SEQ ID NO: 2) without altering the activity of said ASE, whereas an "essential" amino acid residue is required for ASE activity. Other amino acid residues, however, (e.g., those that are not conserved or only semiconserved in the domain having ASE activity) may not be essential for activity and thus are likely to be amenable to alteration without altering ASE activity.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding ASEs that contain modified amino acid residues that are not essential for ASE activity. Such ASEs differ in amino acid sequence from a sequence contained in



SEQ ID NO: 2 yet retain at least one of the ASE activities described herein. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least  
5 about 50% homologous to an amino acid sequence of SEQ ID NO: 2 and is capable of participation in the metabolism of compounds necessary for the construction of cellular membranes in *Isochrysis galbana*, or in the transport of molecules across these membranes, or has one or more activities set forth in Tab. 2.  
10 Preferably, the protein encoded by the nucleic acid molecule is at least about 50-60% homologous to one of the sequences in SEQUENCE ID NO: 2, more preferably at least about 60-70% homologous to one of the sequences in SEQ ID NO: 2, even more preferably at least about 70-80%, 80-90%, 90-95% homologous to  
15 one of the sequences in SEQUENCE ID NO: 2, and most preferably at least about 96%, 97%, 98%, or 99% homologous to one of the sequences in SEQ ID NO: 2.

To determine the percent homology of two amino acid sequences  
20 (e.g., one of the sequences of SEQ ID NO: 2 and a mutant form thereof) or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of one protein or nucleic acid for optimal alignment with the other protein or nucleic acid). The amino acid residues  
25 or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in one sequence (e.g., one of the sequences of SEQ ID NO: 2) is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence (e.g., a mutant form of the form  
30 selected from SEQ ID NO: 2); then the molecules are homologous at that position (i.e., as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity"). The percent homology between the two sequences is a function of the number of identical positions shared by the  
35 sequences (i.e., % homology = numbers of identical positions/total numbers of positions x 100).

An isolated nucleic acid molecule encoding an ASE homologous to a protein sequence of SEQ ID NO: 2 can be created by introducing  
40 one or more nucleotide substitutions, additions or deletions into a nucleotide sequence of SEQ ID NO: 1 such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into one of the sequences of SEQ ID NO: 1 or its derivatives by standard  
45 techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted nonessential amino acid

residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. 5 These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, 10 proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in an ASE is preferably replaced with another amino acid residue 15 from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of the ASE coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for an ASE activity described herein to identify mutants that retain 20 ASE activity. Following mutagenesis of one of the sequences of SEQ ID NO: 1, the encoded protein can be expressed recombinantly and the activity of the protein can be determined using, for example, assays described herein (see Examples).

25 A further known technique for directed evolution and mutagenesis of gene sequences encoding enzymes is gene shuffling (Stemmer, PNAS 1994, 91: 10747-10751, WO 97/20078 and WO 98/13487). Gene shuffling is a method for the combination of gene fragments and can be combined with error prone PCR in order to further enhance 30 the genetic variability of resulting sequences and encoded enzymatic activities. A premise for such an approach is a suitable screening system. In the case of elongases high throughput metabolite measurements facilitated by MALDI-TOF, gas chromatography-mass spectroscopy, thin layer chromatography 35 or liquid chromatography-mass spectroscopy or other suitable combinations or methods can be used to monitor the appearance of new compounds or products in the hydrophobic fraction.

In addition to the nucleic acid molecules encoding ASEs described 40 above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense to an isolated nucleic acid comprising a nucleotide sequence which encodes a polypeptide which elongates  $\alpha$ -linolenic acid ( $C_{18:3} \Delta 9, 12, 15$ ) by at least two carbon atoms whereas  $\gamma$ -linolenic acid ( $C_{18:3} \Delta 6, 9, 12$ ) is not 45 elongated, selected from the group consisting of

## 30

- a) a nucleic acid sequence depicted in SEQ ID NO: 1,  
b) a nucleic acid sequence which encodes a polypeptide depicted  
in SEQ ID NO: 2,

5

- c) derivatives of the sequence depicted in SEQ ID NO: 1,  
which encodes polypeptides having at least 50% homology  
to the sequence encoding amino acid sequences depicted  
in SEQ ID NO: 2 and which sequences function as an elongase.

10

An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly,  
15 an antisense nucleic acid can form hydrogen bonds with a sense nucleic acid. The antisense nucleic acid can be complementary to an entire ASE coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence  
20 encoding an ASE. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (e.g., the entire coding region starting with and ending with the stop codon, i.e. the the last codon before the stop codon). In another embodiment, the antisense  
25 nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding ASE. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions). It is also  
30 possible to use the inverted repeat technology combining an antisense fragment with a portion of the antisense fragment in sense orientation linked by either an adapter sequence or an excisable intron (Abstract Book of the 6<sup>th</sup> Intern. Congr. Of Plant Mol Biol. ISPMB, Quebec June 18-24, 2000, Abstract No. S20-9 by  
35 Green et al.).

Given the coding strand sequences encoding the ASE disclosed herein (e.g., the sequences set forth in SEQ ID NO: 1), antisense nucleic acids of the invention can be designed according to the  
40 rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of ASE mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of ASE mRNA. For example, the antisense oligonucleotide can be  
45 complementary to the region surrounding the translation start site of ASE mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 and more

nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl)uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methyl ester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl)uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

35

The antisense nucleic acid molecules of the invention are typically administered to a cell or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an ASE to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. The antisense molecule can be modified such that it specifically binds to a receptor or an antigen expressed on the selected cell surface, e.g., by linking the antisense nucleic

40

45



- acid molecule to a peptide or an antibody which binds to a cell surface receptor or antigen. The antisense nucleic acid molecule can also be delivered to cells using the vectors described below. To achieve sufficient intracellular concentrations of
- 5 the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong prokaryotic, viral, or eukaryotic promoter, including plant promoters, are preferred.
- 10 In another embodiment, the antisense nucleic acid molecule of the invention is an anomeric nucleic acid molecule. An anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual units, the strands run parallel to each other (Gaultier et al. (1987)
- 15 Nucleic Acids. Res. 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-O-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res. 15:6131-6148) or a chimeric RNA-DNA analog (Inoue et al. (1987) FEBS Lett. 215:327-330).
- 20 In another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead
- 25 ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave ASE mRNA transcripts to thereby inhibit translation of ASE mRNA. A ribozyme having specificity for an ASE-encoding nucleic acid can be designed on the basis of the nucleotide sequence of an ASE
- 30 cDNA disclosed herein in SEQ ID NO: 1 or on the basis of a heterologous sequence to be isolated according to methods taught in this invention. For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to
- 35 be cleaved in an ASE-encoding mRNA. See, e.g., Cech et al. US 4,987,071 and Cech et al. US 5,116,742. Alternatively, ASE mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel, D. and Szostak, J.W. (1993) Science 261:1411-1418.
- 40
- Alternatively, ASE gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of an ASE nucleotide sequence (e.g., an ASE promoter and/or enhancers) to form triple helical structures that prevent transcription of
- 45 an ASE gene in target cells. See generally Helene, C. (1991) Anticancer Drug Res. 6(6):569-84; Helene, C. et al. (1992) Ann.

N.Y. Acad. Sci. 660:27-36; and Maher, L.J. (1992) Bioassays 14(12):807-15.

#### B. Gene construct

5

Another embodiment of the invention is a novel gene construct comprising an isolated nucleic acid derived from a plant which encodes a polypeptide which elongates  $\alpha$ -linolenic acid ( $C_{18:3}$   $\Delta^9, 12, 15$ ) by at least two carbon atoms but not  $\gamma$ -linolenic acid ( $C_{18:3}$   $\Delta^6, 9, 12$ ), or the gene sequence of SEQ ID NO: 1, its homologs, derivatives or analogs as defined above which have been functionally linked to one or more regulatory signals, advantageously to increase gene expression. Examples of these regulatory sequences are sequences to which inducers or repressors bind and thus regulate the expression of the nucleic acid. In addition to these novel regulatory sequences, the natural regulation of these sequences in front of the actual structural genes can still be present and, where appropriate, have been genetically modified so that the natural regulation has been switched off and the expression of the genes has been increased. The gene construct can, however, also have a simpler structure, that is to say no additional regulatory signals have been inserted in front of the sequence SEQ ID NO: 1 or its homologs, and the natural promoter with its regulation has not been deleted. Instead, the natural regulatory sequence has been mutated so that regulation no longer takes place, and gene expression is enhanced. The gene construct may additionally advantageously comprise one or more so-called enhancer sequences functionally linked to the promoter and making increased expression of the nucleic acid sequence possible. It is also possible to insert at the 3' end of the DNA sequences additional advantageous sequences, such as further regulatory elements or terminators. The elongase genes may be present in one or more copies in the gene construct. It is advantageous for further genes to be present in the gene construct for insertion of further genes into organisms.

Advantageous regulatory sequences for the novel process are present, for example, in promoters such as cos-, tac-, trp-, tet-, trp-tet-, lpp-, lac-, lpp-lac-, lacI<sup>q</sup>-, T7-, T5-, T3-, gal-, trc-, ara-, SP6-,  $\lambda$ -P<sub>R</sub>- or  $\lambda$ -P<sub>L</sub>-promoter and are advantageously used in Gram-negative bacteria. Further advantageous regulatory sequences are present, for example, in the Gram-positive promoters amy and SPO2, in the yeast or fungal promoters ADC1, MF $\alpha$ , AC, P-60, CYC1, GAPDH, TEF, rp28, ADH or in the plant promoters CaMV/35S [Franck et al., Cell 21 (1980) 285 - 294], PRP1 [Ward et al., Plant. Mol. Biol. 22 (1993)], SSU, OCS, lib4,

usp, STLS1, B33, nos or in the ubiquitin or phaseolin promoter. Also advantageous in this connection are inducible promoters such as the promoters described in EP-A-0 388 186 (benzenesulfonamide-inducible), Plant J. 2, 1992: 397 - 404 (Gatz et al., tetracyclin-inducible), EP-A-0 335 528 (abscisic acid-inducible) or WO 93/21334 (ethanol or cyclohexenol inducible). Additional useful plant promoters are the cytosolic FBPase promoter or ST-LSI promoter of the potato (Stockhaus et al., EMBO J. 8, 1989, 2445), the phosphoribosylpyrophosphate amidotransferase promoter of Glycine max (gene bank accession No. U87999) or the node-specific promoter described in EP-A-0 249 676. Particularly advantageous promoters are promoters which allow the expression in tissues which are involved in the fatty acid biosynthesis. Most particularly advantageous are seed specific promoters such as usp-, LEB4-, phaseolin or napin promoter. Additional particularly advantageous promoters are seed-specific promoters which can be used for monocots or dicots and which are described in US 5,608,152 (napin promoter from rapeseed), WO 98/45461 (phaseolin promoter from Arabidopsis), US 5,504,200 (phaseolin promoter from Phaseolus vulgaris), WO 91/13980 (Bce4 promoter from Brassica), Baeumlein et al., Plant J., 2, 2, 1992: 233-239 (LEB4 promoter from legumes); said promoters are useful in dicots. The following promoters are useful for example in monocotyledons: lpt-2- or lpt-1-promoter from barley (WO 95/15389 and WO 95/23230), hordein promoter from barley and other useful promoters described in WO 99/16890.

It is possible in principle to use all natural promoters with their regulatory sequences like those mentioned above for the novel process. It is also possible and advantageous in addition to use synthetic promoters.

The gene construct may, as described above, also comprise further genes which are to be inserted into the organisms. It is possible and advantageous to insert and express in host organisms regulatory genes such as genes for inducers, repressors or enzymes which intervene by their enzymatic activity in the regulation, or one or more or all genes of a biosynthetic pathway. These genes can be heterologous or homologous in origin. The inserted genes may have their own promoter or else be under the control of the promoter of the sequence SEQ ID NO: 1 or its homologs.

The gene construct advantageously comprises, for expression of the other genes present, additional 3' and/or 5' terminal regulatory sequences to enhance expression, which are selected

for optimal expression depending on the selected host organism and gene or genes.

These regulatory sequences are intended to make specific  
5 expression of the genes and protein expression possible as mentioned above. This may mean, depending on the host organism, for example that the gene is expressed or overexpressed only after induction, or that it is immediately expressed and/or overexpressed.

10

The regulatory sequences or factors may moreover preferably have a beneficial effect on expression of the introduced genes, and thus increase it. It is possible in this way for the regulatory elements to be enhanced advantageously at the transcription level  
15 by using strong transcription signals such as promoters and/or enhancers. However, in addition, it is also possible to enhance translation by, for example, improving the stability of the mRNA.

In addition the inventive gene construct preferably comprises  
20 additional genes of different biochemical pathways, for example genes for the synthesis of vitamins, carotinoids, sugars such as monosaccharides, oligosaccharides or polysaccharides, or fatty acid biosynthesis genes, more preferably the gene construct comprises fatty acid biosynthesis genes such as desaturases,  
25 hydroxylases, Acyl-ACP-thioesterases, elongases, acetylenases, synthetases or reductases such as  $\Delta 19$ -,  $\Delta 17$ -,  $\Delta 15$ -,  $\Delta 12$ -,  $\Delta 9$ -,  $\Delta 8$ -,  $\Delta 6$ -,  $\Delta 5$ -,  $\Delta 4$ -desaturases, hydroxylases, elongases,  $\Delta 12$ -acetylenase, Acyl-ACP-thioesterases,  $\beta$ -ketoacyl-ACP-synthases or  $\beta$ -ketoacyl-ACP-reductases. Preferably the gene construct  
30 comprises fatty acid biosynthesis genes selected from the group consisting of  $\Delta 19$ -,  $\Delta 17$ -,  $\Delta 15$ -,  $\Delta 12$ -,  $\Delta 9$ -,  $\Delta 8$ -,  $\Delta 6$ -,  $\Delta 5$ -,  $\Delta 4$ -desaturases, hydroxylases, elongases,  $\Delta 12$ -acetylenase, acyl-ACP-thioesterases,  $\beta$ -ketoacyl-ACP-synthases or  $\beta$ -ketoacyl-ACP-reductases.

35

### C. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an ASE (or  
40 a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type  
45 of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they have



been introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., nonepisomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise at least one inventive nucleic acid or at least one inventive gene construct of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells used for expression, which is or are linked operably to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "linked operably" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence and these sequences are fused to each other so that both sequences fulfill the proposed function ascribed to the sequence used (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990) or see: Gruber and Crosby, in: Methods in Plant Molecular Biology and Biotechnology, CRC Press, Boca Raton, Florida, eds.: Glick and Thompson, Chapter 7, 89-108 including the references therein. Regulatory sequences include those which govern constitutive expression of a nucleotide sequence in many types of host cell and those which govern direct expression of the nucleotide sequence only in certain host cells under certain conditions. It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be

introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., ASEs, mutant forms of ASEs, fusion proteins, etc.).

5

The recombinant expression vectors of the invention can be designed for expression of ASEs in prokaryotic or eukaryotic cells. For example, ASE genes can be expressed in bacterial cells such as *C. glutamicum*, insect cells (using baculovirus expression  
10 vectors), yeast and other fungal cells (see Romanos, M.A. et al. (1992) Foreign gene expression in yeast: a review, *Yeast* 8: 423-488; van den Hondel, C.A.M.J.J. et al. (1991) Heterologous gene expression in filamentous fungi, in: *More Gene Manipulations in Fungi*, J.W. Bennet & L.L. Lasure, eds., p. 396-428: Academic  
15 Press: San Diego; and van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) Gene transfer systems and vector development for filamentous fungi, in: *Applied Molecular Genetics of Fungi*, Peberdy, J.F. et al., eds., p. 1-28, Cambridge University Press: Cambridge), algae (Falciatore et al., 1999, *Marine*  
20 *Biotechnology*.1, 3:239-251), ciliates of the types: Holotrichia, Peritrichia, Spirotrichia, Suctorina, Tetrahymena, Paramecium, Colpidium, Glaucoma, Platyophrya, Potomacus, Pseudocohnilembus, Euplotes, Engelmaniella, and Stylonychia, especially of the genus *Stylonychia lemnae*, with vectors following a transformation  
25 method as described in WO9801572, and multicellular plant cells (see Schmidt, R. and Willmitzer, L. (1988), High efficiency *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* leaf and cotyledon explants, *Plant Cell Rep.*: 583-586); *Plant Molecular Biology and Biotechnology*, C Press, Boca Raton,  
30 Florida, chapter 6/7, p. 71-119 (1993); F.F. White, B. Jenes et al., *Techniques for Gene Transfer*, in: *Transgenic Plants*, Vol. 1, Engineering and Utilization, eds.: Kung und R. Wu, Academic Press (1993), 128-43; Potrykus, *Annu. Rev. Plant Physiol. Plant Molec. Biol.* 42 (1991), 205-225 (and references cited therein) or  
35 mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated in vitro, for example using T7 promoter regulatory sequences and  
40 T7 polymerase.

Expression of proteins in prokaryotes is most often carried out with vectors containing constitutive or inducible promoters directing the expression of either fusion or nonfusion proteins.  
45 Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein but also to the C-terminus or fused within suitable regions in

## 38

the proteins. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in  
5 affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their  
10 cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL  
15 (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. In one embodiment, the coding sequence of the elongase ASE is cloned into a pGEX expression  
20 vector to create a vector encoding a fusion protein comprising, from the N-terminus to the C-terminus, GST thrombin cleavage site X protein. The fusion protein can be purified by affinity chromatography using glutathione-agarose resin. Recombinant ASE unfused to GST can be recovered by cleavage of the fusion protein  
25 with thrombin.

Examples of suitable inducible nonfusion *E. coli* expression vectors include pTrc (Amann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *Gene Expression Technology: Methods in*  
30 *Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gnl0-lac fusion promoter mediated by a  
35 coexpressed viral RNA polymerase (T7 gnl). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident  $\lambda$  prophage harboring a T7 gnl gene under the transcriptional control of the lacUV 5 promoter.

40 Other vectors which are useful in prokaryotic organisms are known to a person skilled in the art; such vectors are for example in *E. coli* pLG338, pACYC184, the pBR series such as pBR322, the pUC series such as pUC18 or pUC19, the M13mp series, pKC30, pRep4, pHS1, pHS2, pPLC236, pMBL24, pLG200, pUR290, pIN-III<sup>113</sup>-B1,  $\lambda$ gt11  
45 or pBdCI, in *Streptomyces* pIJ101, pIJ364, pIJ702 or pIJ361, in

Bacillus pUB110, pC194 or pBD214, in Corynebacterium pSA77 or pAJ667.

One strategy to maximize recombinant protein expression is to  
5 express the protein in a host bacteria with an impaired capacity  
to proteolytically cleave the recombinant protein (Gottesman, S.,  
Gene Expression Technology: Methods in Enzymology 185, Academic  
Press, San Diego, California (1990) 119-128). Another strategy  
is to alter the nucleic acid sequence of the nucleic acid to be  
10 inserted into an expression vector so that the individual codons  
for each amino acid are those preferentially utilized in the  
bacterium chosen for expression, such as *C. glutamicum* (Wada  
et al. (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration  
of nucleic acid sequences of the invention can be carried out  
15 by standard DNA synthesis techniques.

In another embodiment, the ASE expression vector is a yeast  
expression vector. Examples of vectors for expression in yeast  
*S. cerevisiae* include pYepSec1 (Baldari, et al., (1987) *Embo J.*  
20 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943),  
pJRY88 (Schultz et al., (1987) *Gene* 54:113-123), and pYES2  
(Invitrogen Corporation, San Diego, CA). Vectors and methods for  
the construction of vectors appropriate for use in other fungi,  
such as the filamentous fungi, include those detailed in: van den  
25 Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and  
vector development for filamentous fungi, in: *Applied Molecular  
Genetics of Fungi*, J.F. Peberdy, et al., eds., p. 1-28, Cambridge  
University Press: Cambridge or in: *More Gene Manipulations in  
Fungi* [J.W. Bennet & L.L. Lasure, eds., p. 396-428: Academic  
30 Press: San Diego]. Additional useful yeast vectors are for  
example 2 $\mu$ M, pAG-1, YEpl3 or pEMBLYe23.

Alternatively, the ASEs of the invention can be expressed in  
insect cells using baculovirus expression vectors. Baculovirus  
35 vectors available for expression of proteins in cultured insect  
cells (e.g., Sf9 cells) include the pAc series (Smith et al.  
(1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow  
and Summers (1989) *Virology* 170:31-39).

40 The abovementioned vectors are only a small overview of possible  
useful vectors. Additional plasmids are well known by the skilled  
artisan and are described for example in: *Cloning Vectors* (Eds.  
Pouwels P.H. et al. Elsevier, Amsterdam-New York-Oxford, 1985,  
ISBN 0 444 904018).



In another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman et al. (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyomavirus, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Nonlimiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al. (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji et al. (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Patent Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

In another embodiment, the ASEs of the invention may be expressed in unicellular plant cells (such as algae), see Falciatore et al., 1999, *Marine Biotechnology*. 1 (3):239-251 and references therein, and plant cells from higher plants (e.g., the spermatophytes, such as crop plants). Examples of plant expression vectors include those detailed in: Becker, D., Kemper, E., Schell, J. and Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20: 1195-1197; and Bevan, M.W. (1984) "Binary *Agrobacterium* vectors for plant transformation, *Nucl. Acid. Res.* 12: 8711-8721; *Vectors for Gene Transfer in Higher Plants*; in:

## 41

Transgenic Plants, Vol. 1, Engineering and Utilization, eds.: Kung und R. Wu, Academic Press, 1993, p. 15-38.

A plant expression cassette preferably contains regulatory  
5 sequences capable of driving gene expression in plants cells  
and which are linked operably so that each sequence can fulfill  
its function such as termination of transcription, such as  
polyadenylation signals. Preferred polyadenylation signals are  
those originating from *Agrobacterium tumefaciens* T-DNA such as  
10 the gene 3 known as octopine synthase of the Ti-plasmid pTiACH5  
(Gielen et al., EMBO J. 3 (1984), 835 ff) or functional  
equivalents thereof but also all other terminators functionally  
active in plants are suitable.

15 As plant gene expression is very often not limited on  
transcriptional levels a plant expression cassette preferably  
contains other operably linked sequences like translational  
enhancers such as the overdrive sequence containing the  
5'-untranslated leader sequence from tobacco mosaic virus  
20 enhancing the protein per RNA ratio (Gallie et al 1987, Nucl.  
Acids Research 15:8693-8711).

Plant gene expression has to be linked operably to an appropriate  
promoter conferring gene expression in a time, cell or tissue  
25 specific manner. Preferred are promoters driving constitutive  
expression (Benfey et al., EMBO J. 8 (1989) 2195-2202) like those  
derived from plant viruses like the 35S CaMV (Franck et al.,  
Cell 21 1980) 285-294), the 19S CaMV (see also US 5,352,605 and  
WO 84/02913) or plant promoters like those from Rubisco small  
30 subunit described in US 4,962,028. Additionally vATPase-gene  
promoters such as a 1153 basepair fragment from *Beta vulgaris*  
(Plant Mol Biol, 1999, 39:463-475) can be used to drive ASE gene  
expression alone or in combination with other PUFA biosynthesis  
genes.

35 Other preferred sequences for use in operable linkage in plant  
gene expression cassettes are targeting sequences necessary to  
direct the gene product in its appropriate cell compartment (for  
review see Kermode, Crit. Rev. Plant Sci. 15, 4 (1996), 285-423  
40 and references cited therein) such as the vacuole, the nucleus,  
all types of plastids like amyloplasts, chloroplasts,  
chromoplasts, the extracellular space, mitochondria, the  
endoplasmic reticulum, oil bodies, peroxisomes and other  
compartments of plant cells.

Plant gene expression can also be facilitated via a chemically inducible promoter (for review see Gatz 1997, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 48:89-108). Chemically inducible promoters are especially suitable if gene expression is wanted to occur in a time specific manner. Examples of such promoters are a salicylic acid inducible promoter (WO 95/19443), a tetracycline inducible promoter (Gatz et al., (1992) *Plant J.* 2, 397-404) and an ethanol inducible promoter (WO 93/21334).

Also promoters responding to biotic or abiotic stress conditions are suitable promoters such as the pathogen inducible PRP1 gene promoter (Ward et al., *Plant. Mol. Biol.* 22 (1993), 361-366), the heat inducible hsp80 promoter from tomato (US 5,187,267), cold inducible alpha-amylase promoter from potato (WO 96/12814) or the wound-inducible pinII promoter (EP-A-0 375 091).

Especially those promoters are preferred which confer gene expression in tissues and organs where lipid and oil biosynthesis occurs, in seed cells such as cells of the endosperm and the developing embryo. Suitable promoters are the napin-gene promoter from rapeseed (US 5,608,152), the USP-promoter from *Vicia faba* (Baeumlein et al., *Mol Gen Genet*, 1991, 225 (3):459-67), the oleosin-promoter from *Arabidopsis* (WO 98/45461), the phaseolin-promoter from *Phaseolus vulgaris* (US 5,504,200), the Bce4-promoter from *Brassica* (WO 91/13980) or the legumin B4 promoter (LeB4; Baeumlein et al., 1992, *Plant Journal*, 2 (2):233-9) as well as promoters conferring seed specific expression in monocot plants like maize, barley, wheat, rye, rice etc. Suitable promoters to note are the lpt2 or lpt1-gene promoter from barley (WO 95/15389 and WO 95/23230) or those described in WO 99/16890 (promoters from the barley hordein-gene, the rice glutelin gene, the rice oryzin gene, the rice prolamin gene, the wheat gliadin gene, wheat glutelin gene, the maize zein gene, the oat glutelin gene, the *Sorghum kasirin*-gene, the rye secalin gene).

Also especially suited are promoters that confer plastid-specific gene expression as plastids are the compartment where precursors and some end products of lipid biosynthesis are synthesized. Suitable promoters such as the viral RNA-polymerase promoter are described in WO 95/16783 and WO 97/06250 and the clpP-promoter from *Arabidopsis* described in WO 99/46394.

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a

manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to PSE mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the  
5 continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of  
10 a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using  
15 antisense genes see Weintraub, H. et al., Antisense RNA as a molecular tool for genetic analysis, Reviews - Trends in Genetics, Vol. 1(1) 1986 and Mol et al., 1990, FEBS Letters 268:427-430.

20 Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the  
25 progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

30 A host cell can be any prokaryotic or eukaryotic cell. For example, an ASE can be expressed in bacterial cells such as *C. glutamicum*, insect cells, fungal cells or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells), algae,  
35 ciliates, plant cells, fungi or other microorganisms like *C. glutamicum*. Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic  
40 cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection", conjugation and transduction are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate  
45 or calcium chloride coprecipitation, DEAE-dextran-mediated transfection, lipofection, natural competence, chemical-mediated transfer, or electroporation. Suitable methods for transforming



or transfecting host cells including plant cells can be found in Sambrook, et al. (Molecular Cloning: A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) and other laboratory manuals 5 such as Methods in Molecular Biology, 1995, Vol. 44, Agrobacterium protocols, ed: Gartland and Davey, Humana Press, Totowa, New Jersey.

For stable transfection of mammalian cells, it is known that, 10 depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host 15 cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as hygromycin and methotrexate or, in plants, those which confer resistance towards a herbicide such as imidazolinones, sulfonylurea, glyphosate or glufosinate. Nucleic acid encoding a 20 selectable marker can be introduced into a host cell on the same vector as that encoding an ASE or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by, for example, drug selection (e.g., cells that have incorporated the selectable marker gene will survive, 25 while the other cells die).

To create a homologously-recombinant microorganism, a vector is prepared which contains at least a portion of an ASE gene into which a deletion, addition or substitution has been introduced 30 to thereby alter, e.g., functionally disrupt, the ASE gene. Preferably, this ASE gene is an *Isochrysis galbana* ASE gene, but it can be a homolog from a related plant or even from an alga, mammalian, yeast, or insect source. In a preferred embodiment, the vector is designed such that, upon homologous recombination, 35 the endogenous ASE gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a knock-out vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous ASE gene is mutated or otherwise altered but still encodes a functional 40 protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous ASE). To create a point mutation via homologous recombination also DNA-RNA hybrids can be used known as chimeraplasty known from Cole-Strauss et al. 1999, Nucleic Acids Research 27(5):1323-1330 and Kmiec, Gene 45 therapy. 1999, American Scientist. 87(3):240-247.

In the homologously-recombinant vector, the altered portion of the ASE gene is flanked at its 5' and 3' ends by additional nucleic acid of the ASE gene to allow for homologous recombination to occur between the exogenous ASE gene carried by a vector and an endogenous ASE gene in a microorganism or plant. The additional flanking ASE nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several hundreds of basepairs up to kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see e.g., Thomas, K.R., and Capecchi, M.R. (1987) Cell 51: 503 for a description of homologous recombination vectors or Strepp et al., 1998, PNAS, 95 (8):4368-4373 for cDNA based recombination in *Isochrysis galbana*). The vector is introduced into a microorganism or plant cell (e.g., via *Agrobacterium* mediated gene transfer, biolistic transformation, polyethylene glycol or other applicable methods) and cells in which the introduced ASE gene has homologously recombined with the endogenous ASE gene are selected, using art-known techniques. In the case of plant cells the AHAS gene described in Ott et al., J. Mol. Biol. 1996, 263:359-360 is especially suitable for marker gene expression and resistance towards imidazolinone or sulfonylurea type herbicides.

In another embodiment, recombinant organisms such as microorganisms can be produced which contain selected systems which allow for regulated expression of the introduced gene. For example, inclusion of an ASE gene in a vector placing it under control of the Lac operon permits expression of the ASE gene only in the presence of IPTG. Such regulatory systems are well known in the art. Recombinant organisms means an organism which comprises an inventive nucleic acid sequence, a gene construct or a vector in the cell or inside the genome at a place which is not the "natural" place or at the "natural" place but modified in a manner which is not the natural manner; that means the coding sequence is modified and/or the regulatory sequence is modified. Modified means single nucleotides or one or more codons are changed in comparison to the natural sequence, preferably one or more codons, more preferably one to six codons.

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) an ASE. An alternate method can be applied in addition in plants by the direct transfer of DNA into developing flowers via electroporation or *Agrobacterium*-mediated gene transfer. Accordingly, the invention further provides methods for producing ASEs using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention

(into which a recombinant expression vector encoding an ASE has been introduced, or into whose genome has been introduced a gene encoding a wild-type or altered ASE) in a suitable medium until ASE is produced. In another embodiment, the method further  
5 comprises isolating ASEs from the medium or the host cell.

Host cells suitable in principle to take up the nucleic acid of the invention, the novel gene construct or the inventive vector are all prokaryotic or eukaryotic organisms. The host organisms  
10 advantageously used are organisms such as bacteria, fungi, yeasts, animal or plant cells. Additional advantageous organisms are animals or preferably plants or parts thereof. Fungi, yeasts or plants are preferably used, particularly preferably fungi or plants, very particularly preferably plants such as oilseed  
15 plants containing high amounts of lipid compounds such as rapeseed, evening primrose, canola, peanut, linseed, soybean, safflower, sunflower, borage or plants such as maize, wheat, rye, oat, triticale, rice, barley, cotton, manihot, pepper, tagetes, solanaceous plants such as potato, tobacco, eggplant, and tomato,  
20 Vicia species, pea, alfalfa, bushy plants (coffee, cacao, tea), Salix species, trees (oil palm, coconut) and perennial grasses and forage crops. Particularly preferred plants of the invention are oilseed plants such as soybean, peanut, rapeseed, canola, sunflower, safflower, trees (oil palm, coconut).

25

#### D. Isolated ASE

Another aspect of the invention pertains to isolated ASEs, and biologically active portions thereof. An "isolated" or "purified"  
30 protein or biologically active portion thereof is substantially free of cellular material when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of ASE in which the  
35 protein is separated from cellular components of the cells in which it is naturally or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of ASE having less than about 30% (by dry weight) of non-ASE (also referred to herein as a  
40 "contaminating protein"), more preferably less than about 20% of non-ASE, still more preferably less than about 10% of non-ASE, and most preferably less than about 5% of non-ASE. When the ASE or biologically active portion thereof is recombinantly produced, it is also substantially free of culture medium, i.e., culture  
45 medium represents less than 20%, more preferably less than 10%, and most preferably less than about 5% of the volume of the protein preparation. The language "substantially free of chemical

precursors or other chemicals" includes preparations of ASE in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of ASE having less than about 30% (by dry weight) of chemical precursors or non-ASE chemicals, more preferably less than about 20% chemical precursors or non-ASE chemicals, still more preferably less than about 10% chemical precursors or non-ASE chemicals, and most preferably less than about 5% chemical precursors or non-ASE chemicals. In preferred embodiments, isolated proteins or biologically active portions thereof lack contaminating proteins from the same organism from which the ASE is derived. Typically, such proteins are produced by recombinant expression of, for example, an *Isochrysis galbana* ASE in other plants than *Isochrysis galbana* or microorganisms such as *C. glutamicum* or ciliates, algae or fungi.

An isolated ASE of the invention or a portion thereof can participate in the metabolism of compounds involved in the construction of cellular membranes in *Isochrysis galbana*, or in the transport of molecules across these membranes, or has one or more of the activities set forth in Tab. 2. In preferred embodiments, the protein or portion thereof comprises an amino acid sequence which is sufficiently homologous to an amino acid sequence of SEQ ID NO: 2 such that the protein or portion thereof maintains the ability to participate in the metabolism of compounds necessary for the construction of cellular membranes in *Isochrysis galbana*, or in the transport of molecules across these membranes. The portion of the protein is preferably a biologically active portion as described herein. In another preferred embodiment, an ASE of the invention has an amino acid sequence shown in SEQ ID NO: 2. In another preferred embodiment, the ASE has an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to a nucleotide sequence of SEQ ID NO: 1. In another preferred embodiment, the ASE has an amino acid sequence which is encoded by a nucleotide sequence that is at least about 50-60%, preferably at least about 60-70%, more preferably at least about 70-80%, 80-90%, 90-95%, and even more preferably at least about 96%, 97%, 98%, 99% or more homologous to one of the amino acid sequences of SEQ ID NO: 2. The preferred ASEs of the present invention also preferably possess at least one of the ASE activities described herein. For example, a preferred ASE of the present invention includes an amino acid sequence encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to a nucleotide sequence



of SEQ ID NO: 1, and which can participate in the metabolism of compounds necessary for the construction of cellular membranes in *Isochrysis galbana*, or in the transport of molecules across these membranes, or which has one or more of the activities set forth in Tab. 2.

In other embodiments, the ASE is substantially homologous to an amino acid sequence of SEQ ID NO: 2 and retains the functional activity of the protein of one of the sequences of SEQ ID NO: 2 yet differs in amino acid sequence due to natural variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the ASE is a protein which comprises an amino acid sequence which is at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80, 80-90, 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to an entire amino acid sequence of SEQUENCE ID NO: 2 and which has at least one of the ASE activities described herein. In another embodiment, the invention pertains to a full *Isochrysis galbana* protein which is substantially homologous to an entire amino acid sequence of SEQ ID NO: 2.

Biologically active portions of an ASE include peptides comprising amino acid sequences derived from the amino acid sequence of an ASE, e.g., the amino acid sequence shown in SEQ ID NO: 2 or the amino acid sequence of a protein homologous to an ASE, which include fewer amino acids than a full length ASE or the full length protein which is homologous to an ASE, and exhibit at least one activity of an ASE. Typically, biologically active portions (peptides, e.g., peptides which are, for example, 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) comprise a domain or motif with at least one activity of an ASE. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the activities described herein. Preferably, the biologically active portions of an ASE include one or more selected domains/motifs or portions thereof having biological activity.

ASEs are preferably produced by recombinant DNA techniques. For example, a nucleic acid molecule encoding the protein is cloned into an expression vector (as described above), the expression vector is introduced into a host cell (as described above) and the ASE is expressed in the host cell. The ASE can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Alternatively to recombinant expression, an ASE, polypeptide, or peptide can be synthesized chemically using standard peptide synthesis

techniques. Moreover, native ASE can be isolated from cells (e.g., endothelial cells), for example using an anti-ASE antibody, which can be produced by standard techniques utilizing an ASE of this invention or fragment thereof.

5

The invention also provides ASE chimeric or fusion proteins. As used herein, an ASE "chimeric protein" or "fusion protein" comprises an ASE polypeptide operatively linked to a non-ASE polypeptide. An "ASE polypeptide" refers to a polypeptide  
10 having an amino acid sequence corresponding to an ASE, whereas a "non-ASE polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the ASE, e.g., a protein which is different from the ASE and which is derived from the same or  
15 a different organism. Within the fusion protein, the term "operatively linked" is intended to indicate that the ASE polypeptide and the non-ASE polypeptide are fused to each other so that both sequences fulfill the predicted function ascribed to the sequence used. The non-ASE polypeptide can be fused to the  
20 N-terminus or C-terminus of the ASE polypeptide. For example, in one embodiment the fusion protein is a GST-ASE fusion protein in which the ASE sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant ASEs. In another embodiment, the fusion protein  
25 is an ASE containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of an ASE can be increased through use of a heterologous signal sequence.

30 An ASE chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or sticky-ended termini for  
35 ligation, restriction enzyme cleavage to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA  
40 synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Current Protocols in  
45 Molecular Biology, eds. Ausubel et al. John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). An

ASE-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the ASE.

Homologs of the ASE can be generated by mutagenesis, e.g.,  
5 discrete point mutation or truncation of the ASE. As used herein, the term "homolog" refers to a variant form of the ASE which acts as an agonist or antagonist of the activity of the ASE. An agonist of the ASE can retain substantially the same, or a subset, of the biological activities of the ASE. An antagonist  
10 of the ASE can inhibit one or more of the activities of the naturally occurring form of the ASE, by, for example, competitively binding to a downstream or upstream member of the cell membrane component metabolic cascade which includes the ASE, or by binding to an ASE which mediates transport of compounds  
15 across such membranes, thereby preventing translocation from taking place.

In an alternative embodiment, homologs of the ASE can be identified by screening combinatorial libraries of mutants, e.g.,  
20 truncation mutants, of the ASE for ASE agonist or antagonist activity. In one embodiment, a variegated library of ASE variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of ASE variants can be produced by, for example,  
25 enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential ASE sequences is expressible as individual polypeptides, or, alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of ASE sequences therein. There  
30 are a variety of methods which can be used to produce libraries of potential ASE homologs from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a  
35 degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential ASE sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang, S.A. (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura  
40 et al. (1984) Science 198:1056; Ike et al. (1983) Nucleic Acid Res. 11:477.

In addition, libraries of fragments of the ASE can be used to generate a variegated population of ASE fragments for screening  
45 and subsequent selection of homologs of an ASE. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an ASE

## 51

coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked  
5 products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes  
10 of the ASE.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products  
15 having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of ASE homologs. The most widely used techniques, which are amenable to high-throughput analysis, for screening large gene libraries typically include cloning the  
20 gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of the desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive  
25 ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify ASE homologs (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331).

30

In another embodiment, cell based assays can be exploited to analyze a variegated ASE library, using further methods well known in the art.

## 35 E. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologs, fusion proteins, primers, vectors, and host cells described herein can be used in one or more of the following methods: identification  
40 of *Isochrysis galbana* and related organisms; mapping of genomes of organisms related to *Isochrysis galbana*; identification and localization of *Isochrysis galbana* sequences of interest; evolutionary studies; determination of ASE regions required for function; modulation of an ASE activity; modulation of the  
45 metabolism of one or more cell membrane components; modulation of the transmembrane transport of one or more compounds; and



modulation of cellular production of a desired compound, such as a fine chemical, including PUFAs.

The ASE nucleic acid molecules of the invention have a variety of  
5 uses. First, they may be used to identify an organism as being  
Isochrysis galbana or a close relative thereof. Also, they may be  
used to identify the presence of Isochrysis galbana or a relative  
thereof in a mixed population of microorganisms. The invention  
provides the nucleic acid sequences of a number of Isochrysis  
10 galbana genes; by probing the extracted genomic DNA of a culture  
of a unique or mixed population of microorganisms under stringent  
conditions with a probe spanning a region of an Isochrysis  
galbana gene which is unique to this organism, one can ascertain  
whether this organism is present. Although Isochrysis galbana  
15 itself is not used for the commercial production of  
polyunsaturated acids, algae are the only known plants beside  
mosses that produce more than a few percent of their total lipids  
as PUFAs. Therefore DNA sequences related to ASEs are especially  
suited to be used for PUFA production in other organisms.

20 Further, the nucleic acid molecules and protein molecules of the  
invention may serve as markers for specific regions of the  
genome. This has utility not only in the mapping of the genome,  
but also for functional studies of Isochrysis galbana proteins.  
25 For example, to identify the region of the genome to which a  
particular Isochrysis galbana DNA-binding protein binds, the  
Isochrysis galbana genome could be digested, and the fragments  
incubated with the DNA-binding protein. Those which bind the  
protein may be additionally probed with the nucleic acid  
30 molecules of the invention, preferably with readily detectable  
labels; binding of such a nucleic acid molecule to the genome  
fragment enables the localization of the fragment on the genome  
map of Isochrysis galbana, and, when performed multiple times  
with different enzymes, facilitates a rapid determination of the  
35 nucleic acid sequence to which the protein binds. Further, the  
nucleic acid molecules of the invention may be sufficiently  
homologous to the sequences of related species such that these  
nucleic acid molecules may serve as markers for the construction  
of a genomic map for related algae, such as Isochrysis galbana.

40 The ASE nucleic acid molecules of the invention are also useful  
for evolutionary and protein structural studies. The metabolic  
and transport processes in which the molecules of the invention  
participate are utilized by a wide variety of prokaryotic and  
45 eukaryotic cells; by comparing the sequences of the nucleic acid  
molecules of the present invention to those encoding similar  
enzymes from other organisms, the evolutionary relatedness of the

organisms can be assessed. Similarly, such a comparison permits an assessment of which regions of the sequence are conserved and which are not, which may aid in determining those regions of the protein which are essential for the functioning of the enzyme.

- 5 This type of determination is of value for protein engineering studies and may give an indication of what the protein can tolerate in terms of mutagenesis without losing function.

Manipulation of the ASE nucleic acid molecules of the invention  
10 may result in the production of ASEs having functional differences from the wild-type ASEs. These proteins may be improved in efficiency or activity, may be present in greater numbers in the cell than is usual, or may be decreased in efficiency or activity. Improved efficiency or activity means for  
15 example the enzyme has a higher selectivity and/or activity than the original enzyme, preferably at least 10% higher, particularly preferably at least 20% higher activity, most particularly preferably at least 30% higher activity.

- 20 There are a number of mechanisms by which the alteration of an ASE of the invention may directly affect the yield, production, and/or efficiency of production of a fine chemical incorporating such an altered protein. Recovery of fine chemical compounds from large-scale cultures of ciliates, algae or fungi is significantly  
25 improved if the cell secretes the desired compounds, since such compounds may be readily isolated from the culture medium (as opposed to extracted from the mass of cultured cells). Otherwise purification can be improved if preferably the cell stores compounds in a specialized compartment having a kind of  
30 concentrating mechanism in vivo. In the case of plants expressing ASEs increased transport can lead to improved partitioning within the plant tissue and organs. By either increasing the number or the activity of transporter molecules which export fine chemicals from the cell, it may be possible to increase the amount of the  
35 produced fine chemical which is present in the extracellular medium, thus permitting greater ease of harvesting and purification or, in the case of plants, more efficient partitioning. Conversely, in order to efficiently overproduce one or more fine chemicals, increased amounts of the cofactors,  
40 precursor molecules, and intermediate compounds for the appropriate biosynthetic pathways are required. By increasing the number and/or activity of transporter proteins involved in the import of nutrients, such as carbon sources (i.e., sugars), nitrogen sources (i.e., amino acids, ammonium salts), phosphate,  
45 and sulfur, it may be possible to improve the production of a fine chemical, due to the removal of any nutrient supply limitations on the biosynthetic process. Fatty acids such as

PUFAs and lipids containing PUFAs are themselves desirable fine chemicals, so by optimizing the activity or increasing the number of one or more ASEs of the invention which participate in the biosynthesis of these compounds, or by impairing the activity of one or more genes which are involved in the degradation of these compounds, it may be possible to increase the yield, production, and/or efficiency of production of fatty acid and lipid molecules in ciliates, algae, plants, fungi, yeasts or other microorganisms.

10

The engineering of one or more ASE genes of the invention may also result in ASEs having altered activities which indirectly impact the production of one or more desired fine chemicals from algae, plants, ciliates or fungi. For example, the normal chemical processes of metabolism result in the production of a variety of waste products (e.g., hydrogen peroxide and other reactive oxygen species) which may actively interfere with these same metabolic processes (for example, peroxynitrite is known to nitrate tyrosine side chains, thereby inactivating some enzymes having tyrosine in the active site (Groves, J.T. (1999) *Curr. Opin. Chem. Biol.* 3(2): 226-235)). While these waste products are typically excreted, cells utilized for large-scale fermentative production are optimized for the overproduction of one or more fine chemicals, and thus may produce more waste products than is typical for a wild-type cell. By optimizing the activity of one or more ASEs of the invention, it may be possible to improve the viability of the cell and to maintain efficient metabolic activity, thereby improving the production of the desired product such as PUFAs. Also, the presence of high intracellular levels of the desired fine chemical may actually be toxic to the cell, so by increasing the ability of the cell to secrete these compounds, one may further improve the viability of the cell.

Further, the ASEs of the invention may be manipulated such that the relative amounts of various lipid and fatty acid molecules are altered. This may have a profound effect on the lipid composition of the membrane of the cell. Since each type of lipid has different physical properties, an alteration in the lipid composition of a membrane may significantly alter membrane fluidity. Changes in membrane fluidity can impact the transport of molecules across the membrane, which, as previously explicated, may modify the export of waste products or the produced fine chemical or the import of necessary nutrients. Such membrane fluidity changes may also profoundly affect the integrity of the cell; cells with relatively weaker membranes are more susceptible to abiotic and biotic stress conditions which may damage or kill the cell. By manipulating ASEs involved

## 55

- in the production of fatty acids and lipids for membrane construction such that the resulting membrane has a membrane composition more amenable to the environmental conditions extant in the cultures utilized to produce fine chemicals, a greater proportion of the cells should survive and multiply. Greater numbers of producing cells should translate into greater yields, production, or efficiency of production of the fine chemical from the culture.
- 10 The aforementioned mutagenesis strategies for ASEs to result in increased yields of a fine chemical are not meant to be limiting; variations of these strategies will be readily apparent to one skilled in the art. Using such strategies, and incorporating the mechanisms disclosed herein, the nucleic acid molecules and protein molecules of the invention may be utilized to generate algae, ciliates, plants, animals, fungi or other microorganisms like *C. glutamicum* expressing mutated ASE nucleic acid and protein molecules such that the yield, production, and/or efficiency of production of a desired compound is improved. This desired compound may be any natural product of algae, ciliates, plants, animals or fungi, which includes the final products of biosynthesis pathways and intermediates of naturally-occurring metabolic pathways, as well as molecules which do not naturally occur in the metabolism of said cells, but which are produced by the cells of the invention.

Another embodiment of the invention is a method for production of PUFAs, said method comprising growing an organism which comprises a nucleic acid of the invention, a gene construct of the invention or a vector of the invention which encodes a polypeptide which elongates  $\alpha$ -linolenic acid ( $C_{18:3} \Delta 9, 12, 15$ ) by at least two carbon atoms but not  $\gamma$ -linolenic acid ( $C_{18:3} \Delta 6, 9, 12$ ), under conditions whereby PUFAs are produced in said organism. Preferably the method comprises the growing of an organism which comprises a nucleotide sequence which encodes a polypeptide which elongates  $\alpha$ -linolenic acid ( $C_{18:3} \Delta 9, 12, 15$ ) by at least two carbon atoms whereas  $\gamma$ -linolenic acid ( $C_{18:3} \Delta 6, 9, 12$ ) is not elongated, selected from the group consisting of

- a) a nucleic acid sequence depicted in SEQ ID NO: 1,
- b) a nucleic acid sequence which encodes a polypeptide depicted in SEQ ID NO: 2,



- c) derivatives of the sequence depicted in SEQ ID NO: 1, which encodes polypeptides having at least 50 % homology to the sequence encoding amino acid sequences depicted in SEQ ID NO: 2 and which sequences function as an elongase.

5

More preferably the nucleic acid sequence is derived from a plant, preferably from the genus *Isochrysis*. The used sequence codes for a polypeptide which elongates  $\Delta 9$  fatty acids.

- 10 The PUFAs produced by this method are preferably C<sub>20</sub> or C<sub>22</sub> fatty acid molecules having at least two double bonds in the fatty acid molecule, preferably at least three double bonds.

- Organisms which are useful in the inventive method for the
- 15 production of PUFAs are microorganism such as bacteria like Gram-positive or Gram-negative bacteria or preferably blue algae, ciliates such as *Colpidium* or *Stylonichia*, fungi such as *Mortierella* or *Thraustochytrium* or *Schizochytrium*, algae such as *Phaeodactylum*, and/or plants like maize, wheat, rye, oats,
- 20 triticale, rice, barley, soybean, peanut, cotton, Brassica species like rapeseed, canola and turnip rape, linseed, pepper, sunflower, borage, evening primrose and tagetes, solanaceous plants like potato, tobacco, eggplant, and tomato, Vicia species, pea, manihot, alfalfa, bushy plants (coffee, cacao, tea), *Salix*
- 25 species, trees (oil palm, coconut) and perennial grasses and forage crops, either directly, e.g., mosses or other plants where overexpression or optimization of a fatty acid biosynthesis protein has a direct impact on the yield, production, and/or efficiency of production of the fatty acid from modified
- 30 organisms.

- PUFAs can be produced in the inventive process in the form of an oil, lipid or free fatty acid. PUFAs produced by this method can be isolated by harvesting the organisms either from the culture
- 35 in which they were growing or from the field, disrupting and/or extracting the harvested material with an organic solvent. From said solvent the oil containing lipids, phospholipids, sphingolipids, glycolipids, triacylglycerols and/or free fatty acids with a higher content of PUFAs can be isolated. By basic
- 40 or acid hydrolysis of the lipids, phospholipids, sphingolipids, glycolipids or triacylglycerols, the free fatty acids with a higher content of PUFAs can be isolated. Higher content of PUFAs means at least 1%, preferably 10%, particularly preferably 20%, most particularly preferably 40% more PUFAs than the original
- 45 organism which has no additional nucleic acid coding for the inventive elongase.

Besides the abovementioned methods, plant lipids are extracted preferably from plant material as described by Cahoon et al. (1999) PNAS 96 (22): 12935-12940 and Browse et al. (1986) Analytic Biochemistry 152: 141-145. Qualitative and quantitative  
5 lipid or fatty acid analysis is described in Christie, William W., Advances in Lipid Methodology, Ayr/Scotland : Oily Press. - (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press, 1989 Repr. 1992 - IX,307 p. - (Oily Press Lipid  
10 Library ; 1); "Progress in Lipid Research, Oxford: Pergamon Press, 1(1952) - 16(1977) under the title: Progress in the Chemistry of Fats and Other Lipids.

PUFAs produced by this method are preferably C<sub>20</sub> or C<sub>22</sub> fatty acid  
15 molecules having at least two double bonds in the fatty acid molecule, preferably three to four double bonds, particularly preferably three double bonds. Such C<sub>20</sub> or C<sub>22</sub> fatty acid molecules can be isolated from the organism in the form of an oil, lipid or free fatty acid. Organisms which are useful are for  
20 example the ones mentioned above. Preferred organisms are transgenic plants.

One embodiment of the invention is oils, lipids or fatty acids or fractions thereof produced by the method described above,  
25 particularly preferably an oil, lipid or fatty acid composition comprising PUFAs derived from transgenic plants.

A further embodiment of the invention is the use of said oil, lipid or fatty acid composition in feed, food, cosmetics or  
30 pharmaceuticals.

An additional embodiment of the invention is a monoclonal or polyclonal antibody which specifically interacts with a polypeptide encoded by the inventive nucleic acid sequence  
35 described above and which is produced by a method known by the skilled worker.

A further embodiment of the invention is a kit comprising an inventive nucleotide sequence, a gene construct as claimed, a  
40 vector as claimed or an antibody as described above. Said kit is useful for example for the identification of the protein, the nucleic acid sequence.

The aforementioned mutagenesis strategies for ASEs to result in  
45 increased yields of a fine chemical are not meant to be limiting; variations on these strategies will be readily apparent to one skilled in the art. Using such strategies, and incorporating

the mechanisms disclosed herein, the nucleic acid molecules and protein molecules of the invention may be utilized to generate algae, ciliates, plants, fungi or other microorganisms like *C. glutamicum* expressing mutated ASE nucleic acid and protein molecules such that the yield, production, and/or efficiency of production of a desired compound is improved. This desired compound may be any natural product of algae, ciliates, plants, fungi or *C. glutamicum*, which includes the final products of biosynthesis pathways and intermediates of naturally-occurring metabolic pathways, as well as molecules which do not naturally occur in the metabolism of said cells, but which are produced by said cells of the invention.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patent applications, patents, and published patent applications cited throughout this application are hereby incorporated by reference.

## 20 Examples

### Example 1: General processes

#### a) Cloning processes and general methods

Cloning processes such as, for example, restriction cleavages, agarose gel electrophoresis, purification of DNA fragments, transfer of nucleic acids to nitrocellulose and nylon membranes, linkage of DNA fragments, transformation of *Escherichia coli* and yeast cells, growth of bacteria and sequence analysis of recombinant DNA were carried out as described in Sambrook et al. (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) or Kaiser, Michaelis and Mitchell (1994) "Methods in Yeast Genetics" (Cold Spring Harbor Laboratory Press: ISBN 0-87969-451-3). Transformation and cultivation of algae such as *Chlorella* or *Phaeodactylum* are performed as described by El-Sheekh (1999), *Biologia Plantarum* 42: 209-216; Apt et al. (1996), *Molecular and General Genetics* 252 (5): 872-9.

#### 40 b) Chemicals:

The chemicals used were obtained, if not mentioned otherwise in the text, in p.a. quality from the companies Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) and Sigma (Deisenhofen). Solutions were prepared using purified, pyrogen-free water, designated as H<sub>2</sub>O in the following text, from a Milli-Q water system water purification plant (Millipore,

Eschborn). Restriction endonucleases, DNA-modifying enzymes and molecular biology kits were obtained from the companies AGS (Heidelberg), Amersham (Braunschweig), Biometra (Göttingen), Boehringer (Mannheim), Genomed (Bad Oeynhausen), New England Biolabs (Schwalbach/Taunus), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Pharmacia (Freiburg), Qiagen (Hilden) and Stratagene (Amsterdam, Netherlands). They were used, if not mentioned otherwise, according to the manufacturer's instructions.

10

c) Algal material

For this study, algae of the species *Isochrysis galbana* CCAP 927/1 were used, obtained from the Culture Collection of Algae and Protozoa, Centre for Coastal and Marine Sciences, Dunstaffnage Marine Laboratory, Oban, Argyll; UK.

Cultivation of algae

20 *Isochrysis galbana* was cultured using the f/2 medium containing 10% organic medium as described by Guillard, R.R.L. [1975; Culture of phytoplankton for feeding marine invertebrates. In: Smith, W.L. and Chanley, M.H. (Eds.) Culture of marine Invertebrate animals, NY Plenum Press, pp. 29-60.]. *Isochrysis* 25 *galbana* was cultured at 14°C under continuous light and a light intensity of 30 microEinstein in glass vessels with shaking at 100 rpm.

The f/2 medium consists of:

30

995.5 ml artificial sea water containing:

1 ml NaNO<sub>3</sub> (75 g/l),  
1 ml NaH<sub>2</sub>PO<sub>4</sub> (5 g/l),  
35 1 ml trace element solution,  
1 ml Tris/Cl pH 8.0,  
0.5 ml f/2 vitamin solution

Trace element solution:

40

Na<sub>2</sub>EDTA (4.36 g/l),  
FeCl<sub>3</sub> (3.15 g/l),

Primary trace elements:

45



60

5           CuSO<sub>4</sub> (10 g/l),  
          ZnSO<sub>4</sub> (22 g/l),  
          CoCl<sub>2</sub> (10 g/l),  
          MnCl<sub>2</sub> (18 g/l),  
          NaMoO<sub>4</sub> (6.3 g/l)

f/2 vitamin solution:

10           biotin: 10 mg/l,  
          thiamine 200 mg/l,  
          vit B12 0.1 mg/l

Org. medium:

15           Na acetate (1 g/l),  
          glucose (6 g/l),  
          Na succinate (3 g/l),  
          Bacto-Tryptone (4 g/l),  
          Yeast extract (2 g/l)

20

Example 2: DNA isolation from algae

The details for the isolation of total DNA relate to the working  
up of one gram fresh weight of material harvested by filtration.

25

CTAB buffer:

30           2% (w/v) N-cetyl-N,N,N-trimethylammonium bromide (CTAB);  
          100 mM Tris HCl pH 8.0;  
          1.4 M NaCl;  
          20 mM EDTA.

N-Laurylsarcosine buffer:

35           10% (w/v) N-laurylsarcosine;  
          100 mM Tris HCl pH 8.0;  
          20 mM EDTA.

40           The material was homogenized under liquid nitrogen with quartz  
          sand in a mortar to give a fine powder and transferred to 2 ml  
          Eppendorf cups. The frozen material was then covered with a layer  
          of 1 ml of decomposition buffer (1 ml CTAB buffer, 100 ml of  
          N-laurylsarcosine buffer, 20 ml of beta-mercaptoethanol and 10 ml  
          of proteinase K solution, 10 mg/ml) and incubated at 60°C for  
45           one hour with continuous shaking. The homogenate obtained was  
          distributed into two Eppendorf vessels (2 ml) and extracted twice  
          by shaking with the same volume of chloroform/isoamyl alcohol.

## 61

(24:1). For phase separation, centrifugation was carried out at 8000 x g and room temperature for 15 min in each case. The DNA was then precipitated at 70°C for 30 min. The precipitated DNA was sedimented at 4°C and 10,000 g for 30 min and resuspended in 5 100 microliters of TE buffer (Sambrook et al., 1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6). For further purification, the DNA was treated with NaCl (1.2 M final concentration) and precipitated again at 70°C for 30 min using twice the volume of absolute ethanol. After a washing step with 10 70% ethanol, the DNA was dried and subsequently taken up in 50 microliters of H<sub>2</sub>O + DNase free RNase (50 mg/ml final concentration). The DNA was dissolved overnight at 4°C and the RNase digestion was subsequently carried out at 37°C for 1 h. Storage of the DNA took place at 4°C.

15

### Example 3: Isolation of total RNA and poly(A)+ RNA from algae

For the investigation of transcripts, both total RNA and poly(A)+ RNA were isolated.

20

Algal cultures were harvested by centrifugation at 3000 g for 5 minutes. The pellets were immediately frozen in liquid nitrogen (-70°C). Algal material (1 g) was homogenized with a pestle in a mortar under liquid nitrogen. The material was desintegrated to 25 homogeneity in two volumes of buffer which was TriPure<sup>®</sup> Isolation Reagent (Roche). The total RNA was then isolated following the manufacturer's protocol.

Isolation of Poly(A)+ RNA was carried out using Amersham 30 Pharmacia mRNA Isolation Kit following the instructions of the manufacturer's protocol.

After determination of the concentration of the RNA or of the poly(A)+ RNA, the RNA was precipitated by addition of 1/10 35 volumes of 3 M sodium acetate pH 4.6 and 2 volumes of ethanol and stored at -70°C.

### Example 4: cDNA library construction

40 Double stranded cDNA was synthesised using the cDNA Synthesis Kit from Stratagene following the manufacturer's protocol. It was then passed through a Sephacyl S-400 Spun Column from a cDNA Synthesis Kit (Amersham Pharmacia) to remove adapters and smaller molecules. cDNA eluted from the column was phenol extracted, 45 ethanol precipitated and ligated to the arms of the Uni-Zap vector and packed into  $\lambda$  phages using Ready-To-Go Lambda Packaging Kit (Amersham Pharmacia Biotech) following the manufacturer's

62

instructions. A library of  $1 \times 10^6$  pfu was obtained with the majority of the inserts ranging from 0.4-2 kb.

Example 5: Identification of the ASE1 gene and analysis of the  
5 cDNA-clone Ig\_ASE1

From an alignment of known elongase sequences (from *M. alpina*, *S. cerevisiae* (Elo1, Elo2, Elo3), *C. elegans* (F56H11.4, F41H10.8)) the common motif MYXYYFL was chosen for oligonucleotide design.

10

The reverse complement oligo

5'-A(A/G)(A/G)AA(A/G)TA(A/G)TAIII(G/A)TACAT-3' (I = deoxyinosine)

15 was synthesised and used in touchdown PCR with a universal T3 promoter primer (5'-AATTAACCCCTCACTAAAGGG-3') using an *Isochrysis galbana* cDNA library as template.

The PCR conditions were:

20

94°C for 3 min (1 cycle)

94°C for 15 sec, 52°C for 30 sec, 72°C for 45 sec (4 cycles)

94°C for 15 sec, 52°C for 30 sec (with 1°C decrement every cycle),  
72°C for 45 sec (10 cycles)

25 94°C for 15 sec, 42°C for 30 sec, 72°C for 45 sec (25 cycles)  
72°C for 6 min (1 cycle).

A PCR product of about 650 bp was cloned and sequenced and the deduced amino acid sequence was found to align with the putative  
30 elongase sequence compilation. The gene-specific (sense) primer  
5'-ACTCGAAGCTCTTCACATGG-3'

was synthesised and used in a further library PCR reaction with a universal M13 forward primer  
(5'-GTAAACGACGGCCAGT-3')

35 using the following conditions:

94°C for 3 min (1 cycle)

94°C for 15 sec, 55°C for 30 sec, 72°C for [lacuna] (10 cycles)

94°C for 15 sec, 55°C for 30 sec, 72°C for 1 min 33 sec (with  
40 3 sec increment every cycle) (20 cycles)

72°C for 6 min (1 cycle).

A PCR product of about 850 bp was cloned and sequenced. The two  
PCR product sequences overlapped, confirming that they were  
45 ultimately derived from a single gene.

## 63

Those cDNA clones isolated from cDNA libraries as described in Example 6 were used for DNA sequencing according to standard methods, in particular by the chain termination method using the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit 5 (Perkin-Elmer, Weiterstadt, Germany). Sequencing was carried out subsequent to plasmid recovery from cDNA libraries via in vivo excision and retransformation of DH10B on agar plates (material and protocol details from Stratagene, Amsterdam, Netherlands).

10 Plasmid DNA was prepared from overnight grown *E. coli* cultures grown in Luria-Broth medium containing ampicillin [see Sambrook et al. (1989) (Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6)].

15 Sequencing primers with the following nucleotide sequences were used:

5'-CAGGAAACAGCTATGACC-3'

5'-CTAAAGGGAACAAAAGCTG-3'

20 5'-TGTAACGACGGCCAGT-3'

The complete nucleotide sequence of the cDNA consisted of about 1064 bp. It contained an open reading frame of 789 bp encoding 263 amino acids. The protein sequence shares just low identity or 25 similarity with known genes such elongases which are required for medium-chain-length fatty acid elongation in yeast (Toke & Martin, 1996, Isolation and characterization of a gene affecting fatty acid elongation in *Saccharomyces cerevisiae*. Journal of Biological Chemistry 271,18413-18422.).

30

Table 1. Alignment of *Isochrysis galbana* elongase with homologous sequences.

Gene	M. alpina	Human	Mouse	Yeast	C. elegans
35 Ig_ASE1	27	25.5 (1) 20.2 (2)	24.3	21 (3) 23.2 (4) 23.6 (5)	19

The values in the table are percentage identities from pairwise 40 alignment using DNAMAN (Lynnon Biosoft). Parameters used: Matrix: BLOSUM, Alignment method: optimal K-tuple: 2, Gap open: 10, Gap penalty: 4, Gap extension: 0.1; 1=ELOVL4; 2=Helol; 3=Elo1; 4=Elo2; 5=Elo3.

45



The sequences have been taken from human ELOVL (article 1, sequence 1), human Helol (article 1, sequence 2), M. alpina (Glelo, article 3), C. elegans (article 4), mouse Elovl4 (article 1), yeast (sequence 3, 4, 5 from articles 5 and 6).

5

1. Zhang et al., Nature Gen. 27: 89-93 (2001)
2. Leonard et al., Biochem. J. 350: 765-770 (2000).
- 10 3. Parker-Barnes et al., Proc. Natl. Sci. USA 97: 8284-8289, (2000).
4. Beaudoin et al., Proc Natl. Sci. USA 97: 6421-6426. (2000)
- 15 5. Toke and Martin, J. Biol. Chem. 271: 18413-18422 (1996)
6. Oh et al., J. Biol. Chem. 272: 17373-17384 (1997).

Pairwise alignments of the Ig\_ASE1 gene and Mortierella and mouse  
20 homologs are shown in Figure 1 and Figure 2.

The following parameters were used for the alignments:

Pairwise alignments:	Fixed penalty: 10
25 Ktuple: 1	Floating penalty: 10
Number of diagonals: 3	Window size: 5
Weight matrix (protein): PAM 250	Gap penalty: 5

Example 6: Identification of genes by hybridization

30

Gene sequences can be used to identify homologous or heterologous genes from cDNA or genomic libraries.

Homologous genes (e.g. full length cDNA clones homologous to  
35 [lacuna] and homologs) can be isolated via nucleic acid hybridization using for example cDNA libraries: Depending on the abundance of the gene of interest 100 000 up to 1 000 000 recombinant bacteriophages are plated and transferred to a nylon membrane. After denaturation with alkali, DNA is immobilized on  
40 the membrane by e.g. UV crosslinking. Hybridization is carried out at high stringency conditions. In aqueous solution hybridization and washing are performed at an ionic strength of 1 M NaCl and a temperature of 68°C. Hybridization probes are generated by e.g. radioactive (<sup>32</sup>P) nick transcription labeling  
45 (High Prime, Roche, Mannheim, Germany). Signals are detected by autoradiography.

## 65

Partially homologous or heterologous genes that are related but not identical can be identified analogously to the above described procedure using low stringency hybridization and washing conditions. For aqueous hybridization the ionic strength is normally kept at 1 M NaCl while the temperature is progressively lowered from 68 to 42°C.

Isolation of gene sequences with homologies only in a distinct domain of (for example) 10-20 amino acids can be carried out by using synthetic radiolabeled oligonucleotide probes. Radiolabeled oligonucleotides are prepared by phosphorylation of the 5'-end of two complementary oligonucleotides with T4 polynucleotide kinase. The complementary oligonucleotides are annealed and ligated to form concatemers. The double stranded concatemers are then radiolabeled by for example nick transcription. Hybridization is normally performed at low stringency conditions using high oligonucleotide concentrations.

Oligonucleotide hybridization solution:

- 6x SSC
- 0.01 M sodium phosphate
- 1 mM EDTA (pH 8)
- 0.5 % SDS
- 100 µg/ml denaturated salmon sperm DNA
- 0.1 % nonfat dried milk.

During hybridization temperature is lowered stepwise to 5-10°C below the estimated oligonucleotide T<sub>m</sub> or down to room temperature followed by washing steps and autoradiography. Washing is performed with extremely low stringency such as 3 washing steps using 4x SSC. Further details are described by Sambrook, J. et al. (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press or Ausubel, F.M. et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.

#### Example 7: Plasmids for plant transformation

For plant transformation binary vectors such as pGPTV (Becker et al. 1992, Plant Mol. Biol. 20:1195-1197) or pBinAR can be used (Höfgen and Willmitzer, Plant Science 66 (1990), 221-230). Construction of the binary vectors can be performed by ligation of the cDNA in sense or antisense orientation into the T-DNA.

## 66

5' to the cDNA a plant promotor activates transcription of the cDNA. A polyadenylation sequence is located 3' to the cDNA.

Tissue specific expression can be achieved by using a tissue specific promotor. For example seed specific expression can be achieved by cloning the DC3 or LeB4 or USP promotor 5' to the cDNA. Also any other seed specific promotor element can be used. For constitutive expression within the whole plant the CaMV 35S promotor can be used.

10

The expressed protein can be targeted to a cellular compartment using a signal peptide, for example for plastids, mitochondria or endoplasmatic reticulum (Kermode, Crit. Rev. Plant Sci. 15, 4 (1996), 285-423). The signal peptide is cloned 5' in frame to the cDNA to achieve subcellular localization of the fusion protein.

#### Example 8: Transformation of Agrobacterium

Agrobacterium mediated plant transformation can be performed using for example Agrobacterium strain C58C1 pGV2260 (Deblaere et al. 1984, Nucl. Acids Res. 13, 4777-4788) or GV3101(pMP90) (Koncz and Schell, Mol. Gen. Genet. 204 (1986), 383-396) or LBA4404 (Clontech). Transformation can be performed by standard transformation techniques (Deblaere et al., Nucl. Acids. Res. 13 (1984), 4777-4788).

#### Example 9: Plant transformation

Agrobacterium mediated plant transformation can be performed using standard transformation and regeneration techniques (Gelvin, Stanton B.; Schilperoort, Robert A, Plant Molecular Biology Manual, 2nd Ed. - Dordrecht: Kluwer Academic Publ., 1995. ISBN 0-7923-2731-4; Glick, Bernard R.; Thompson, John E., Methods in Plant Molecular Biology and Biotechnology, Boca Raton: CRC Press, 1993. - 360 pp., ISBN 0-8493-5164-2).

For example rapeseed can be transformed via cotyledon or hypocotyl transformation (Moloney et al., Plant cell Report 8 (1989), 238-242; De Block et al., Plant Physiol. 91 (1989, 694-701). Use of antibiotics for agrobacterium and plant selection depends on the binary vector and the agrobacterium strain used for transformation. Rapeseed selection is normally performed using kanamycin as selectable plant marker.

45

Agrobacterium mediated gene transfer to flax can be performed using for example a technique described by Mlynarova et al. (1994), Plant Cell Report 13: 282-285.

- 5 Transformation of soybean can be performed using for example a technique described in EP 0424 047, US 322 783 (Pioneer Hi-Bred International) or in EP 0397 687, US 5 376 543, US 5 169 770 (University Toledo).
- 10 Plant transformation using particle bombardment, polyethylene glycol mediated DNA uptake or via the silicon carbide fiber technique is for example described by Freeling and Walbot in: "The maize handbook" (1993) ISBN 3-540-97826-7, Springer Verlag New York.

15

Example 10: *In vivo* mutagenesis

- In vivo mutagenesis of microorganisms can be performed by passage of plasmid (or other vector) DNA through E. coli
- 20 or other microorganisms (e.g. Bacillus spp. or yeasts such as Saccharomyces cerevisiae) which are impaired in their capabilities to maintain the integrity of their genetic information. Typical mutator strains have mutations in the genes for the DNA repair system (e.g., mutHLS, mutD, mutT, etc.; for
- 25 reference, see Rupp, W.D. (1996) DNA repair mechanisms, in: Escherichia coli and Salmonella, p. 2277-2294, ASM: Washington). Such strains are well known to those skilled in the art. The use of such strains is illustrated, for example, in Greener, A. and Callahan, M. (1994) Strategies 7: 32-34. Transfer of mutated
- 30 DNA molecules into plants is preferably done after selection and testing in microorganisms. Transgenic plants are generated according to various examples within the Examples section of this document.

- 35 Example 11: Assessment of the expression of a recombinant gene product in a transformed organism

The activity of a recombinant gene product in the transformed host organism has been measured on the transcriptional and/or

40 on the translational level.

- A useful method to ascertain the level of transcription of the gene (an indicator of the amount of mRNA available for translation of the gene product) is to perform a Northern blot
- 45 (for reference see, for example, Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York or within the abovementioned Examples section), in which a primer designed to



bind to the gene of interest is labeled with a detectable tag (usually radioactive or chemiluminescent), such that when the total RNA of a culture of the organism is extracted, run on gel, transferred to a stable matrix and incubated with this probe, 5 the binding and quantity of binding of the probe indicates the presence and also the quantity of mRNA for this gene. This information is evidence of the degree of transcription of the transformed gene. Total cellular RNA can be prepared from cells, tissues or organs by several methods, all well-known in the art, 10 such as that described in Bormann, E.R. et al. (1992) *Mol. Microbiol.* 6: 317-326.

To assess the presence or relative quantity of protein translated from this mRNA, standard techniques, such as a Western blot, may 15 be employed (see, for example, Ausubel et al. (1988) *Current Protocols in Molecular Biology*, Wiley: New York). In this process, total cellular proteins are extracted, separated by gel electrophoresis, transferred to a matrix such as nitrocellulose, and incubated with a probe, such as an antibody, which 20 specifically binds to the desired protein. This probe is generally tagged with a chemiluminescent or colorimetric label which may be readily detected. The presence and quantity of label observed indicates the presence and quantity of the desired mutant protein present in the cell.

25

Example 12: Analysis of impact of recombinant proteins on the production of the desired product

The effect of the genetic modification in plants, fungi, algae, 30 ciliates or [lacuna] on production of a desired compound (such as fatty acids) can be assessed by growing the modified microorganisms or plants under suitable conditions (such as those described above) and analyzing the medium and/or the cellular component for increased production of the desired product (i.e., 35 lipids or a fatty acid). Such analysis techniques are well known to one skilled in the art, and include spectroscopy, thin layer chromatography, staining methods of various kinds, enzymatic and microbiological methods, and analytical chromatography such as high performance liquid chromatography (see, for example, 40 Ullmann, *Encyclopedia of Industrial Chemistry*, vol. A2, p. 89-90 and p. 443-613, VCH: Weinheim (1985); Fallon, A. et al., (1987) *Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology*, vol. 17; Rehm et al. (1993) *Biotechnology*, vol. 3, Chapter III: Product recovery and 45 purification, pages 469-714, VCH: Weinheim; Belter, P.A. et al. (1988) *Bioseparations: downstream processing for biotechnology*, John Wiley and Sons; Kennedy, J.F. and Cabral, J.M.S. (1992)

Recovery processes for biological materials, John Wiley and Sons; Shaeiwitz, J.A. and Henry, J.D. (1988) Biochemical separations, in: Ullmann's Encyclopedia of Industrial Chemistry, vol. B3, Chapter 11, pages 1-27, VCH: Weinheim; and Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications.).

- Besides the abovementioned methods, plant lipids are extracted from plant material as described by Cahoon et al. (1999) PNAS 96 (22): 12935-12940 and Browse et al. (1986) Analytic Biochemistry 152: 141-145. Qualitative and quantitative lipid or fatty acid analysis is described by Christie, William W., Advances in Lipid Methodology, Ayr/Scotland: Oily Press. - (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press, 1989 Repr. 1992. - IX, 307 pages - (Oily Press Lipid Library; 1); "Progress in Lipid Research, Oxford: Pergamon Press, 1(1952) - 16(1977) under the title: Progress in the Chemistry of Fats and Other Lipids.
- 20 In addition to the measurement of the final product of fermentation, it is also possible to analyze other components of the metabolic pathways utilized for the production of the desired compound, such as intermediates and by-products, to determine the overall efficiency of production of the compound. Analysis methods include measurements of nutrient levels in the medium (e.g., sugars, hydrocarbons, nitrogen sources, phosphates, and other ions), measurements of biomass composition and growth, analysis of the production of common metabolites of biosynthetic pathways, and measurement of gases produced during fermentation.
- 30 Standard methods for these measurements are outlined in Applied Microbial Physiology, A Practical Approach, P.M. Rhodes and P.F. Stanbury, eds., IRL Press, p. 103-129; 131-163; and 165-192 (ISBN: 0199635773) and references cited therein.
- 35 One example is the analysis of fatty acids (abbreviations: FAME, fatty acid methyl ester; GC-MS, gas-liquid chromatography-mass spectrometry; TAG, triacylglycerol; TLC, thin-layer chromatography).
- 40 Unequivocal proof for the presence of fatty acid products can be obtained by the analysis of recombinant organisms following standard analytical procedures: GC, GC-MS or TLC as variously described by Christie and references therein (1997, in: Advances on Lipid Methodology- Fourth ed.: Christie, Oily Press, Dundee, 119-169; 1998, gas-chromatography-mass spectrometry methods, Lipids 33:343-353).

Material to be analyzed can be disintegrated via sonification, glass milling, liquid nitrogen and grinding or via other applicable methods. The material has to be centrifuged after disintegration. The sediment is resuspended in Aqua dest, heated  
5 for 10 min at 100°C, cooled on ice and centrifuged followed by extraction in 0,5 M sulfuric acid in methanol containing 2% dimethoxypropane for 1h at 90°C, leading to hydrolyzed oil and lipid compounds resulting in transmethyalted lipids. These fatty acid methyl esters are extracted in petroleum ether and finally  
10 subjected to GC analysis using a capillary column (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 m, 0.32 mm) at a temperature gradient between 170°C and 240°C for 20 min and 5 min at 240°C. The identity of the resulting fatty acid methyl esters has to be defined by the use of standards available from commercial sources  
15 (i.e. Sigma).

In the case of fatty acids where standards are not available molecule identity has to be shown via derivatization and subsequent GC analysis. For example the localization of triple  
20 bond fatty acids has to be shown via GC-MS after derivatization with 4,4-dimethoxyoxazoline derivatives (Christie, 1998, see above).

#### 25 Example 13: Expression products in heterologous microbial systems

Strains, growth conditions and plasmids

Escherichia coli strain XL1. Blue MRF' kan (Stratagene) was used for subcloning the new elongase Ig\_ASE1 from Isochrysis galbana.  
30 For functional expression of this gene we used the Saccharomyces cerevisiae strain INVSc 1 (Invitrogen Co.). E. coli was grown in Luria-Bertani broth (LB, Duchefa, Haarlem, The Netherlands) at 37°C. When necessary, ampicillin (100 mg/liter) was added and 1.5% (w/v) agar (Difco) was included for solid LB media. S.  
35 cerevisiae was grown at 30°C either in YPG-medium or in complete minimal dropout uracil medium (CMdum; see in: Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., Albright, L.B., Coen, D.M., and Varki, A. (1995), Current Protocols in Molecular Biology, John Wiley &  
40 Sons, New York.) containing either 2% (w/v) raffinose or glucose. For solid media 2 % (w/v) Bacto™ agar (Difco) was included. Plasmids used for cloning and expression were pUC 18 (Pharmacia) and pYES2 (Invitrogen Co.).

#### 45 Example 14: Cloning and Expression of an ALA-PUFA specific Elongase (ASE gene) from Isochrysis galbana in yeast

## 71

## a) Cloning procedures

For expression in yeast, the *Isochrysis galbana* gene Ig\_ASE1 was first modified to create restriction sites and the yeast  
5 consensus sequence for highly efficient translation (Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes (*Cell* 44, 283-292.)). A site adjacent to the start codon was introduced. For amplification of the open reading  
10 frame a pair of primers complementary to its 5'- and 3'-end were synthesized.

Forward primer: 5'-GGTACCATGGCCCTCGCAAACGA-3'

15 Reverse primer: 5'-TAGGACATCCACAATCCAT-3'

The PCR reaction was performed with plasmid-DNA as template in a Thermocycler (Biometra) using *Pfu* DNA polymerase (Stratagene) and the following temperature program: 3 min. at 96°C followed by 25  
20 cycles with 30 s at 96°C, 30 s at 55°C and 3 min. at 72°C, 1 cycle with 10 min. at 72°C and stop at 4°C.

The correct size of the amplified DNA fragment of about 800 bp was confirmed by Agarose-TBE gel electrophoresis. The amplified  
25 DNA was extracted from the gel with the QIAquick Gel Extraktion Kit (QIAGEN) and ligated into the T/A-site of the vector pCR 21 (Invitrogen) using the Sure Clone Ligation Kit (Pharmacia). After transformation of *E.coli* XL1 Blue MRF' kan a DNA mini-preparation (Riggs, M.G. & McLachlan, A. (1986), A simplified screening  
30 procedure for large numbers of plasmid mini-preparation. *BioTechniques* 4, 310-313.) was performed with 24 ampicillin-resistant transformants and positive clones were identified by *Bam*HI restriction analysis. The sequence of the cloned PCR product was confirmed by resequencing using the ABI  
35 PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer, Weiterstadt).

The plasmid-DNA of pCR-ASE1 was further restricted with *Kpn*I/*Sac*I and the resulting DNA fragment was ligated into the same  
40 restriction site of the dephosphorylated yeast-*E.coli* shuttle-vector pYES2 resulting in pY2ASE1. After transformation of *E.coli* and DNA mini-preparation from the transformants, the orientation of the DNA fragment within the vector was checked. One clone was grown for DNA maxi-preparation with the Nucleobond®  
45 AX 500 Plasmid-DNA Extraction Kit (Macherey-Nagel, Düringen).



## 72

*Saccharomyces cerevisiae* INVSc1 was transformed with pY2ASE1 and pYES2 by a modified PEG/lithium acetate protocol (Ausubel et al., 1995). After selection on CMDum agar plates containing 2% glucose, four pY2ASE1 transformants and one pYES2 transformant 5 were chosen for further cultivation and functional expression.

## b) Functional expression of elongase activity in yeast

## Preculture:

10

20 ml of CMDum liquid medium containing 2% (w/v) raffinose were inoculated with the transgenic yeast clones (pY2ASE1a-d, pYES2) and grown for 3 days at 30°C, 200 rpm until an optical density at 600 nm (OD<sub>600</sub>) of 1.5-2 was reached.

15

## Main culture:

For expression 20 ml CMDum liquid medium with 2% raffinose and 1% (v/v) Tergitol NP-40 was supplemented with the fatty acid to 20 be tested to a final concentration of 0.003% (w/v). The media were inoculated with the precultures to an OD<sub>600</sub> of 0.05. The expression was induced at an OD<sub>600</sub> of 0.2 with 2% (w/v) galactose for 16 h, after which time the cultures had reached an OD<sub>600</sub> of 0.8-1.2.

25

## c) Fatty Acid Analysis

The total fatty acids were extracted from yeast cultures and analyzed by gas chromatography. For this, cells from 5 ml culture 30 were harvested by centrifugation (1000 x g, 10 min., 4°C) and washed once with 100 mM NaHCO<sub>3</sub>, pH 8.0 to remove residual medium and fatty acids. For preparation of the fatty acid methyl esters (FAMES) the cell pellets were treated with 1 N methanolic H<sub>2</sub>SO<sub>4</sub> and 2% (v/v) dimethoxypropane for 1 h at 80°C. The FAMES were 35 extracted twice with 2 ml petroleum ether, washed once with 100 mM NaHCO<sub>3</sub>, pH 8.0 and once with distilled water and dried with Na<sub>2</sub>SO<sub>4</sub>. The organic solvent was evaporated under a stream of argon and the FAMES were dissolved in 50 µl of petroleum ether. The samples were separated on a ZEBRON ZB-Wax capillary column (30 m, 40 0.32 mm, 0.25 µm; Phenomenex) in a Hewlett Packard 6850 gas chromatograph with a flame ionization detector. The oven temperature was programmed from 70°C (1 min. hold) to 200°C at a rate of 20°C/min., then to 250°C (5 min. hold) at a rate of 5°C/min and finally to 260°C at a rate of 5°C/min. Nitrogen was 45 used as carrier gas (4.5 ml/min. at 70°C). The fatty acids were

identified by comparison with retention times of FAME standards (SIGMA).

- 5 Fatty acid patterns of transgenic yeast strains are shown in Tab. 2

Table 2: Fatty acid patterns in mol% of transgenic yeast strains

10

	- Substrate		+ LA (18:2 n-6)		+ ALA (18:3 n-3)		+GLA (18:3 n-6)	
	+gal	-gal	+gal	-gal	+gal	-gal	+gal	-gal
Induction								
15 16:0	28.7	30.2	27.0	28.9	26.6	28.9	30.0	31.0
16:1 n-9	41.6	42.4	30.7	25.4	30.1	26.4	24.3	24.6
18:0	6.8	6.1	5.7	5.8	6.3	6.3	6.8	6.2
18:1 n-9	22.9	21.3	16.5	13.4	18.4	16.6	14.7	13.4
18:2 n-6*	-	-	11.0	26.5	-	-	-	-
20 18:3 n-6*	-	-	-	-	-	-	24.2	24.8
18:3 n-3*	-	-	-	-	10.2	21.8	-	-
20:2 n-6	-	-	9.1	-	-	-	-	-
20:3 n-3	-	-	-	-	8.4	-	-	-
25 % Elong- ation	0	-	45.3	-	45.2	-	0	-

Explanation to Tab. 2:

Fatty acid elongation of different substrates supplied to transgenic yeast containing pY2ASE1. Exogenous fatty acids  
 30 supplied as substrates for elongation are indicated by an asterisk [\*]. The values given are expressed as mol % of total fatty acid methyl esters identified by GC and FID. In the case of elongated substrates, this is also expressed as a % conversion. Expression of the ASE1 transgene was induced by the addition of  
 35 galactose. Only C18 substrates with a double bond at the Δ9 position were elongated by the ASE1 open reading frame. All values represent the mean of three separate experiments.

GC analysis of FAMES prepared from total lipids of the yeasts  
 40 transformed with pY2ASE1 and grown in the presence of different exogenous fatty acids (ALA, GLA, LA), and their fatty acid patterns are shown in mol% in Table 1. The incorporation of GLA does not yield any elongation product di-homo-GLA (20:3 d8,11,14) while ALA is elongated to yield C20:3 d11,14,17, and LA is  
 45 elongated to yield C20:2 d11, 14.

The transgenic yeast clones transformed with pY2ASE1 and supplied with exogenous substrates show an additional peak in the gas chromatogram (identified by an asterisk [\*] in Figures 3A-D), which has been identified by comparison of retention times 5 as the fed/incorporated fatty acid. A gas chromatography/mass spectroscopy can give additional support to confirm its identity.

Figure 3 A-D shows essentially a GC graph of data presented in Table 2. Explanation to Fig. 3 A-D:

10 GC chromatograms of fatty acid methyl esters extracted from transgenic yeast containing pY2ASE1. Yeast cultures were grown in the presence (indicated by an asterisk) or absence of exogenous fatty acids. Exogenous fatty acids (in the form of sodium salts) were LA (linoleic acid; 18:2  $\Delta^9, 12$ ; 18:2 n-6, see Figure 3B), ALA 15 ( $\alpha$ -linolenic acid; 18:3  $\Delta^9, 12, 15$ ; 18:3 n-3, see Figure 3A), GLA ( $\gamma$ -linolenic acid; 18:3  $\Delta^6, 9, 12$ ; 18:3 n-6, see Figure 3C) or no substrate (Figure 3D). Figure 3B represents expression of the ASE1 ORF as induced by the addition of galactose. After 24 h, yeast cells were harvested by centrifugation, washed to remove 20 exogenous substrate and methylated. Fatty acid methyl esters were separated and detected using standard methods and peaks identified by comigration of known standards. It is clear that Ig\_ASE1 encodes a  $\Delta^9$ -C<sub>18</sub>-PUFA-specific elongating activity. Identified products show that nucleotide sequence of Ig\_ASE1 25 codes for a  $\Delta^9$ -selective C<sub>18</sub> fatty acid elongase from the alga *Isochrysis galbana*, which leads to the formation of new fatty acids in transgenic yeasts.

Further feeding experiments with several other fatty acids can be 30 performed to confirm the substrate selectivity of this elongase in further detail.

Example 15: Purification of the desired product from transformed organisms in general

35 Recovery of the desired product from plant material or fungi, algae, ciliate cells or supernatant of the above-described cultures can be performed by various methods well known in the art. If the desired product is not secreted from the cells, 40 the cells can be harvested from the culture by low-speed centrifugation, the cells can be lysed by standard techniques, such as mechanical force or sonification. Organs of plants can be separated mechanically from other tissue or organs. Following homogenization cellular debris is removed by centrifugation, and 45 the supernatant fraction containing the soluble proteins is retained for further purification of the desired compound. If the product is secreted from desired cells, then the cells are

removed from the culture by low-speed centrifugation, and the supernatant fraction is retained for further purification.

The supernatant fraction from each purification method is  
5 subjected to chromatography with a suitable resin, in which the desired molecule is either retained on the chromatography resin while many of the impurities in the sample are not, or where the impurities are retained by the resin while the sample is not. Such chromatography steps may be repeated as necessary, using  
10 the same or different chromatography resins. One skilled in the art would be well-versed in the selection of appropriate chromatography resins and in their most efficacious application for a particular molecule to be purified. The purified product may be concentrated by filtration or ultrafiltration, and stored  
15 at a temperature at which the stability of the product is maximized.

There are a wide array of purification methods known in the art and the preceding method of purification is not meant to  
20 be limiting. Such purification techniques are described, for example, in Bailey, J.E. & Ollis, D.F. *Biochemical Engineering Fundamentals*, McGraw-Hill: New York (1986).

The identity and purity of the isolated compounds may be assessed by techniques standard in the art. These include high-performance  
25 liquid chromatography (HPLC), spectroscopic methods, staining methods, thin layer chromatography, NIRS, enzymatic assays, or microbiological assays. Such analysis methods are reviewed in: Patek et al. (1994) *Appl. Environ. Microbiol.* 60: 133-140; Malakhova et al. (1996) *Biotekhnologiya* 11: 27-32; and Schmidt et  
30 al. (1998) *Bioprocess Engineer.* 19: 67-70. Ullmann's Encyclopedia of Industrial Chemistry, (1996) vol. A27, VCH: Weinheim, p. 89-90, p. 521-540, p. 540-547, p. 559-566, 575-581 and p. 581-587; Michal, G. (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon,  
35 A. et al. (1987) *Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology*, vol. 17.

#### Equivalents

40 Those skilled in the art will recognize, or will be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed  
45 by the claims.



We claim:

1. An isolated nucleic acid derived from a plant encoding a polypeptide which elongates  $\alpha$ -linolenic acid (C<sub>18:3</sub>  $\Delta$ 9, 12, 15) by at least two carbon atoms whereas  $\gamma$ -linolenic acid (C<sub>18:3</sub>  $\Delta$ 6, 9, 12) is not elongated.
2. An isolated nucleic acid comprising a nucleotide sequence which encodes a polypeptide which elongates  $\alpha$ -linolenic acid (C<sub>18:3</sub>  $\Delta$ 9, 12, 15) by at least two carbon atoms whereas  $\gamma$ -linolenic acid (C<sub>18:3</sub>  $\Delta$ 6, 9, 12) is not elongated, selected from the group consisting of
  - a) a nucleic acid sequence depicted in SEQ ID NO: 1,
  - b) a nucleic acid sequence which encodes a polypeptide depicted in SEQ ID NO: 2,
  - c) derivatives of the sequence depicted in SEQ ID NO: 1, which encodes polypeptides having at least 50 % homology to the sequence encoding amino acid sequences depicted in SEQ ID NO: 2 and which sequences function as an elongase.
3. An isolated nucleic acid sequence as claimed in claim 2, wherein the sequence is derived from a plant.
4. An isolated nucleic acid sequence as claimed in claim 2 or claim 3, wherein the sequence is derived from the genus *Isochrysis*.
5. An isolated nucleic acid sequence as claimed in claim 1 or 2, wherein the polypeptide encoded by the sequence elongates  $\Delta$ 9-fatty acids.
6. An amino acid sequence encoded by an isolated nucleic acid as claimed in any of claims 2 to 4.
7. A gene construct comprising an isolated nucleic acid having the sequence SEQ ID NO: 1 as claimed in any of claims 1 to 4, wherein the nucleic acid is functionally linked to one or more regulatory signals.
8. A gene construct as claimed in claim 7, comprising additional fatty acid biosynthesis genes.

9. A gene construct as claimed in claim 7 or 8, wherein the additional fatty acid biosynthesis genes are selected from the group consisting of  $\Delta 19$ -,  $\Delta 17$ -,  $\Delta 15$ -,  $\Delta 12$ -,  $\Delta 9$ -,  $\Delta 8$ -,  $\Delta 6$ -,  $\Delta 5$ -,  $\Delta 4$ -desaturases, hydroxylases, elongases,  $\Delta 12$ -acetylenase, Acyl-ACP-thioesterasen,  $\beta$ -ketoacyl-ACP-synthases or  $\beta$ -ketoacyl-ACP-reductases.
10. A vector comprising a nucleic acid as claimed in claim 1 or 2 or a gene construct as claimed in any of the claims 7 to 9.
11. An organism comprising at least one nucleic acid as claimed in claim 1 or 2, a gene construct as claimed in any of claims 7 to 9 or a vector as claimed in claim 10.
12. An organism as claimed in claim 11, wherein the organism is a microorganism, an animal or a plant.
13. An organism as claimed in claim 11 or 12, wherein the organism is a transgenic plant.
14. A process for the production of PUFAs, said process comprising growing an organism which comprises a nucleic acid as claimed in claim 1 or 2, a gene construct as claimed in any of the claims 7 to 9 or a vector as claimed in claim 10 which encodes a polypeptide which elongates  $\alpha$ -linolenic acid ( $C_{18:3} \Delta 9, 12, 15$ ) by at least two carbon atoms whereas  $\gamma$ -linolenic acid ( $C_{18:3} \Delta 6, 9, 12$ ) is not elongated, under conditions whereby PUFAs are produced in said organisms.
15. A process as claimed in claim 14, wherein the PUFAs produced by the process are  $C_{20}$  or  $C_{22}$  fatty acid molecules having at least two double bonds in the fatty acid molecule.
16. A process as claimed in claim 14, wherein the  $C_{20}$  or  $C_{22}$  fatty acid molecules are isolated from the organism in the form of an oil, lipid or free fatty acid.
17. A process as claimed in any of claims 14 to 16, wherein the organism is a microorganism, an animal or a plant.
18. A process as claimed in any of claims 14 to 17, wherein the organism is a transgenic plant.
19. A process as claimed in any of the claims 14 to 18, wherein the  $C_{18}$  fatty acid is a fatty acid having three double bonds in the molecule.

## 78

20. An oil, lipid or fatty acid or fraction thereof produced by the process as claimed in any of claims 14 to 19.
21. An oil, lipid or fatty acid composition comprising PUFAs and derived from transgenic plants comprising a nucleotide sequence as claimed in claim 1 or 2, a gene construct as claimed in any of claims 7 to 9 or a vector as claimed in claim 10.
22. The use of an oil, lipid or fatty acid composition as claimed in claim 20 or 21, in feed, food, cosmetics or pharmaceuticals.
23. An antibody which specifically interacts with a polypeptide encoded by a nucleotide sequence as claimed in claim 1 or 2.
24. An antisense nucleotide sequence which is complementary to the nucleotide sequence as claimed in claim 1 or 2.
25. A kit comprising a nucleotide sequence as claimed in claim 1 or 2, a gene construct as claimed in any of claims 7 to 9, a vector as claimed in claim 10 or an antibody as claimed in claim 22.

25

30

35

40

45

Fig. 1 Pairwise alignment of *Isochrysis galbana* elongase (Ig\_ASE1, upper row) with *Mortierella alpina* elongase (M.alpinaGlelo, lower row). Identities are shown in bold characters.

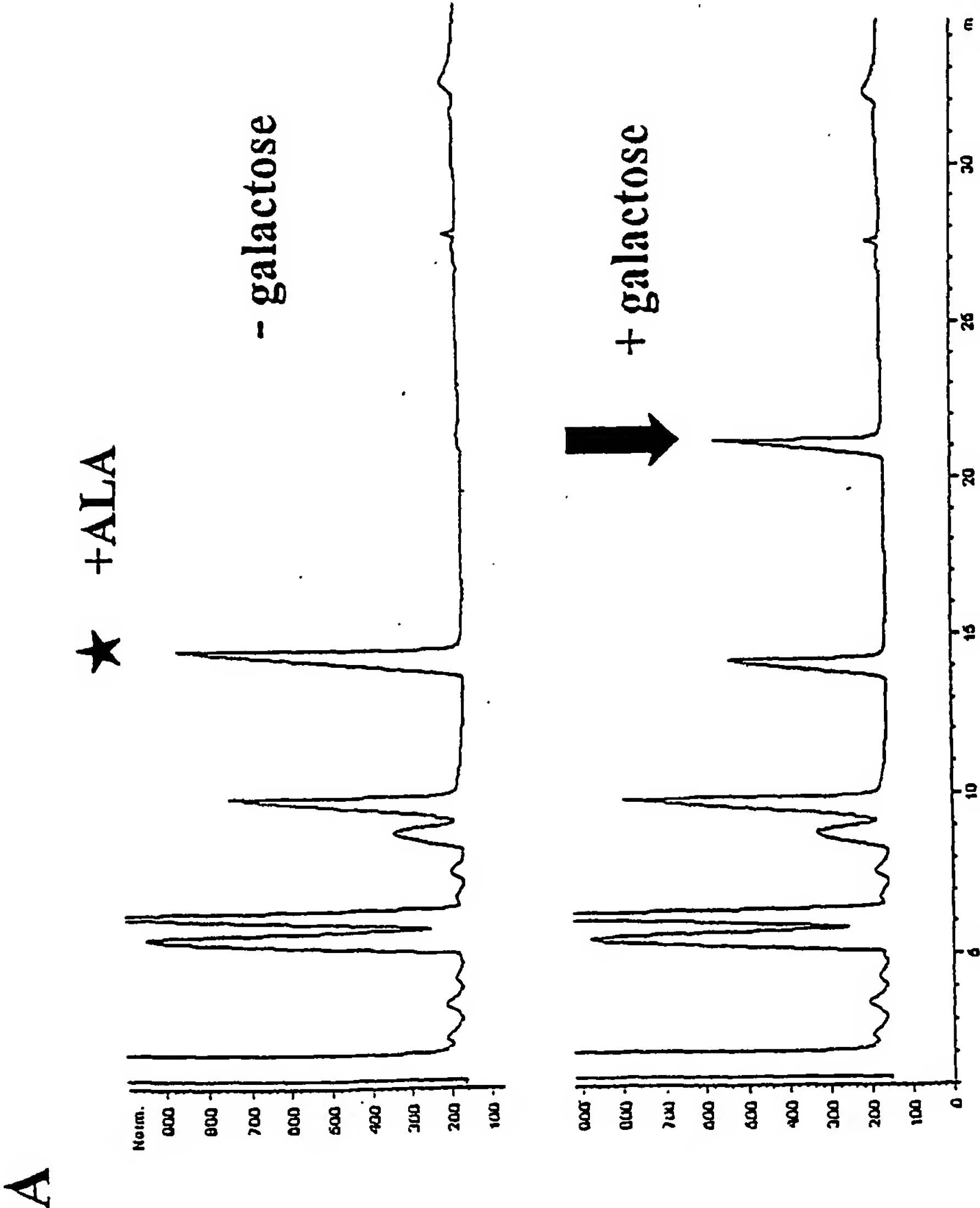
Ig_ASE1	.....MALANDAGERIWAAVTDPEI.....	20
M.alpinaGlelo	MESIAPFLPSKMPQDLFMDLATAIGVRAAPYVDPLEAALVAQAEKYIPTIVHHTR	55
	..LIGTFSYLLLKPLLRNSGLVDEKKGAYRTSMIWNVLLALFSAL.....	64
	GFLVAVESPLARELPLMNPFFHVLLIVLAYLVTVFVGMQIMKNFERFEVKTFSLH	110
	.....SFYVVTATALGWDYGTGAWLRRQTGTPQLFQCPSPVWDSKLEFTWTAK	112
	NFCLVSIAYMCGGILYEAYQANYGLFENAADHTFKGLPMAKMIWL.....	156
	AFYYSKYVEYLDTAWLRV.....SFLQAFHHFGAPWDVYLGIRLHNEGVMIFM	160
	.FYFSKIMEFVDTMIMVLKKNRQISFLHVYHHSSIFTIWLVTFFVAPNGEAYFS	210
	FF.NSFIHTIMYTYGLTAAGYKFA..KPLITAMQICQFVGGFLLVWDYINV.P	211
	AALNSFIHVIIMYGYFFLSALGFKQVSFIKFYITRSQMTQFCMMSVQSSWDMYAMK	265
	CFNSDKGKLFSWAFNYAYVGSVFLFCHFFYQDN.LATKKSAAKAKQL.....	258
	VLGRPGYPFFITALLWFYMTMLGLFYNFYRKNAKLAKAKADAAKEKARKLQ	318



Fig. 2 Pairwise alignment of of polypeptides of Ig\_ASE1 (upper row) and mouse (lower row).  
Identities are shown in bold characters.

IgASE1 Elov14_ (Mus)	.....MALANDAGERIWAAVTDPEILIGTFSY	27
	MGLLDSEPGSVLNAMSTAFNDTVEFYRWTWTIADKRVADWPLMQSPWPTISITL	55
	LLLKPLLNRNSGLVDEKKGAYRTSMIWN...VLLALFSALSFYVTATALGWDYGT	79
	YLLFVWLGPKWMDREPFQMRLVLIYNFGMVLLNLFIFRELFMGSYNAGYSYIC	110
	GAWLRRQTGDTQPPLFQCPSPVWDSKLFWTAKAFYYSKYVEYLDTAWLR.....	129
	QSVDY.....SNDVNEVRIAAALMWYFVSKGVEYLDTVFFILRRKN	151
	..VSFLQAFHHFGAPWDVYLGIRLHNEGVIIFMFF.NSFHTIMYTYGYGLTAAGY	181
	NQVSFLHVYHHC'TMFTLWIGIKWVAGGQAFFGAQMNSFIHVIMYSYGYGLTAFGP	206
	KFKAKPLITAMQICQFVGGFLLVWDYINVPCFNSDKGLFSWAFNAYYVGSVFLL	236
	WIQKYLWWKRYLTMLQLVQFHVTIGHTALSLSLYTDCPPFKMHWALIAAIAISFIFL	261
	FCIFFYQDNLATKKSAGKQL.....	258
	FLNFYTRTYNEPKQS.KTGKTATNGISSNGVKNSEKALENGKPKQNGKPKGE	312

Figure 3 A-D: HPLC chromatograms of fatty acid methyl esters isolated from yeast expressing the Ig\_ASE1 gene product



4/6

Figure 3 A-D: HPLC chromatograms of fatty acid methyl esters isolated from yeast expressing the Ig\_ASE1 gene product

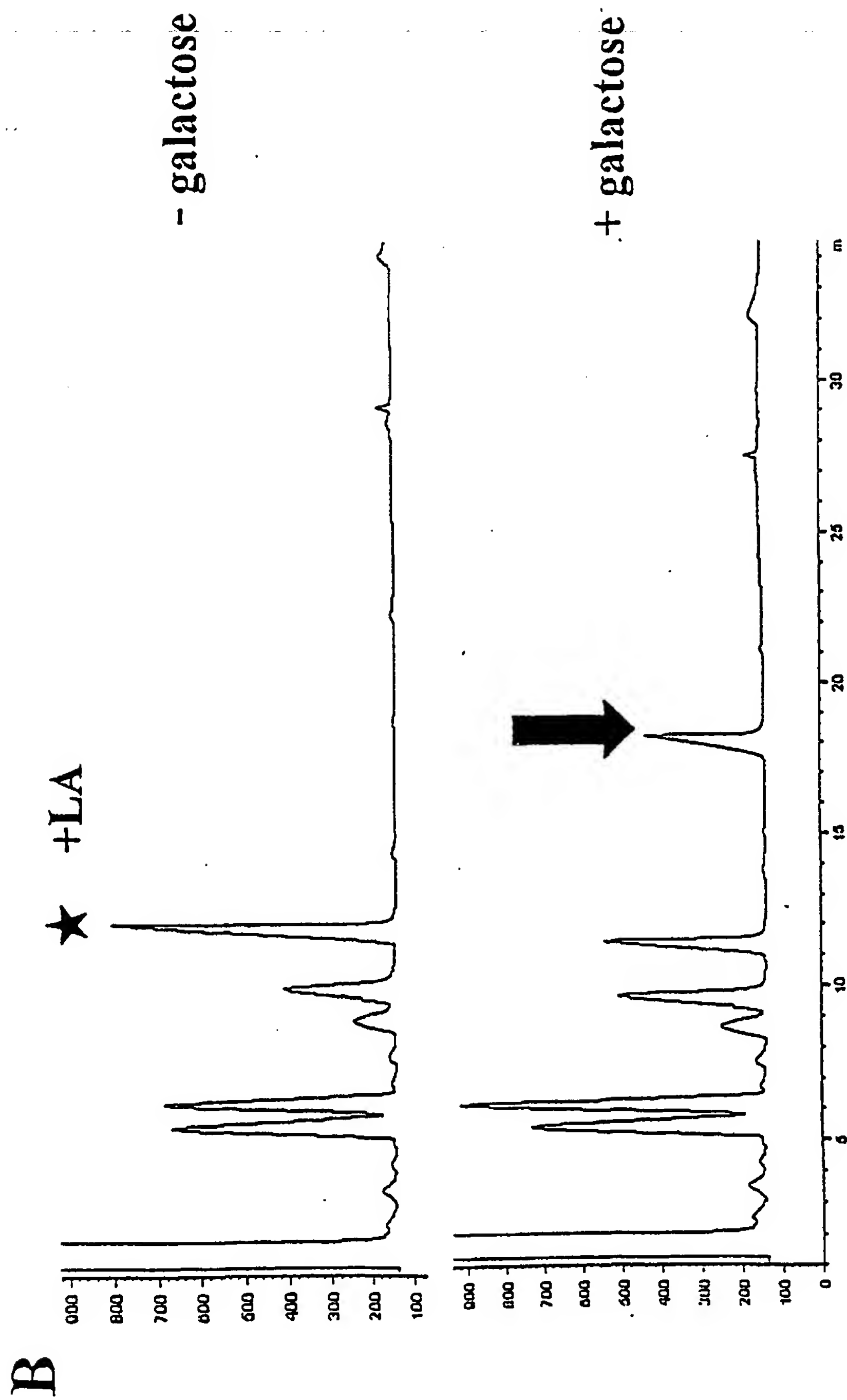


Figure 3 A-D: HPLC chromatograms of fatty acid methyl esters isolated from yeast expressing the Ig\_ASE1 gene product

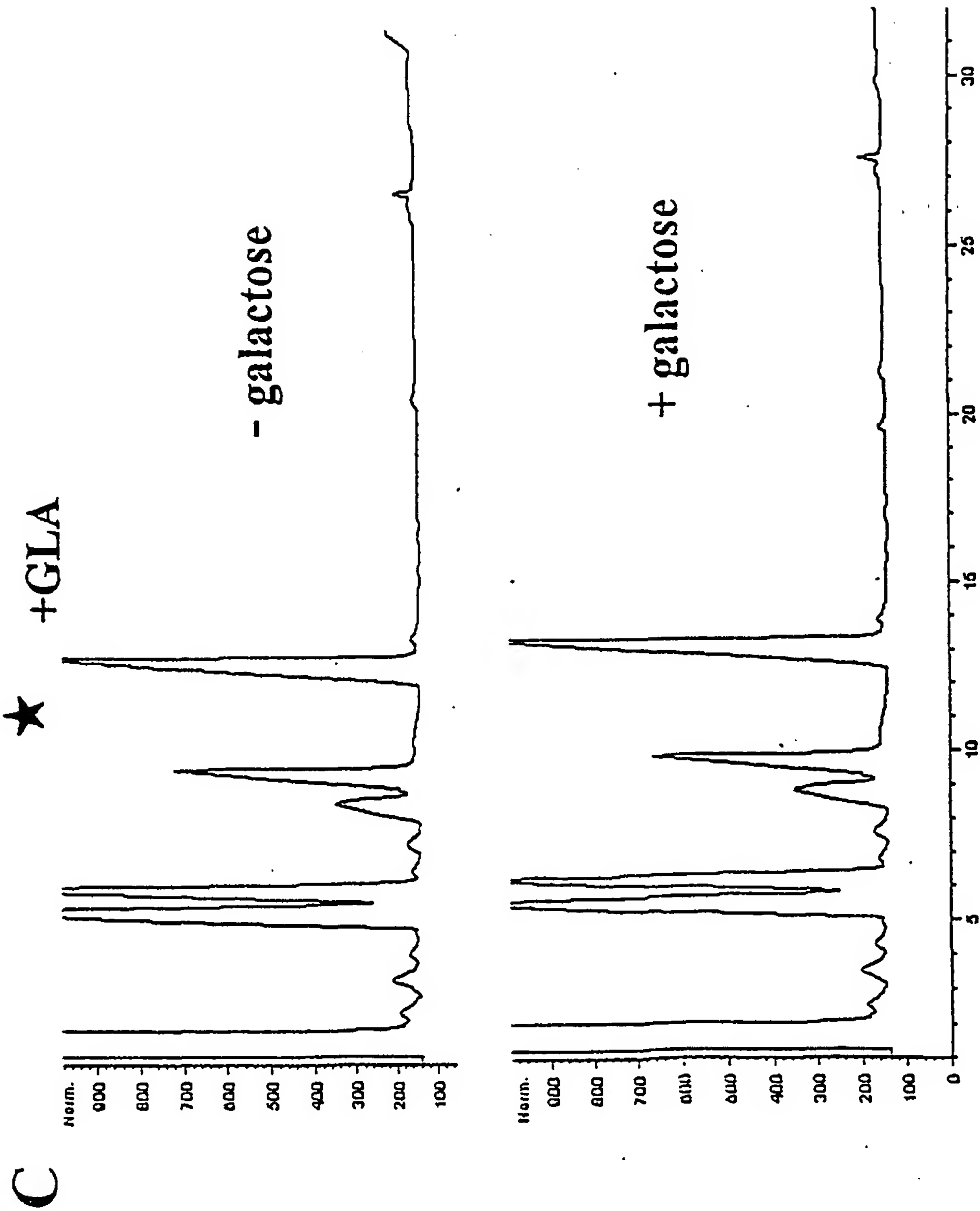
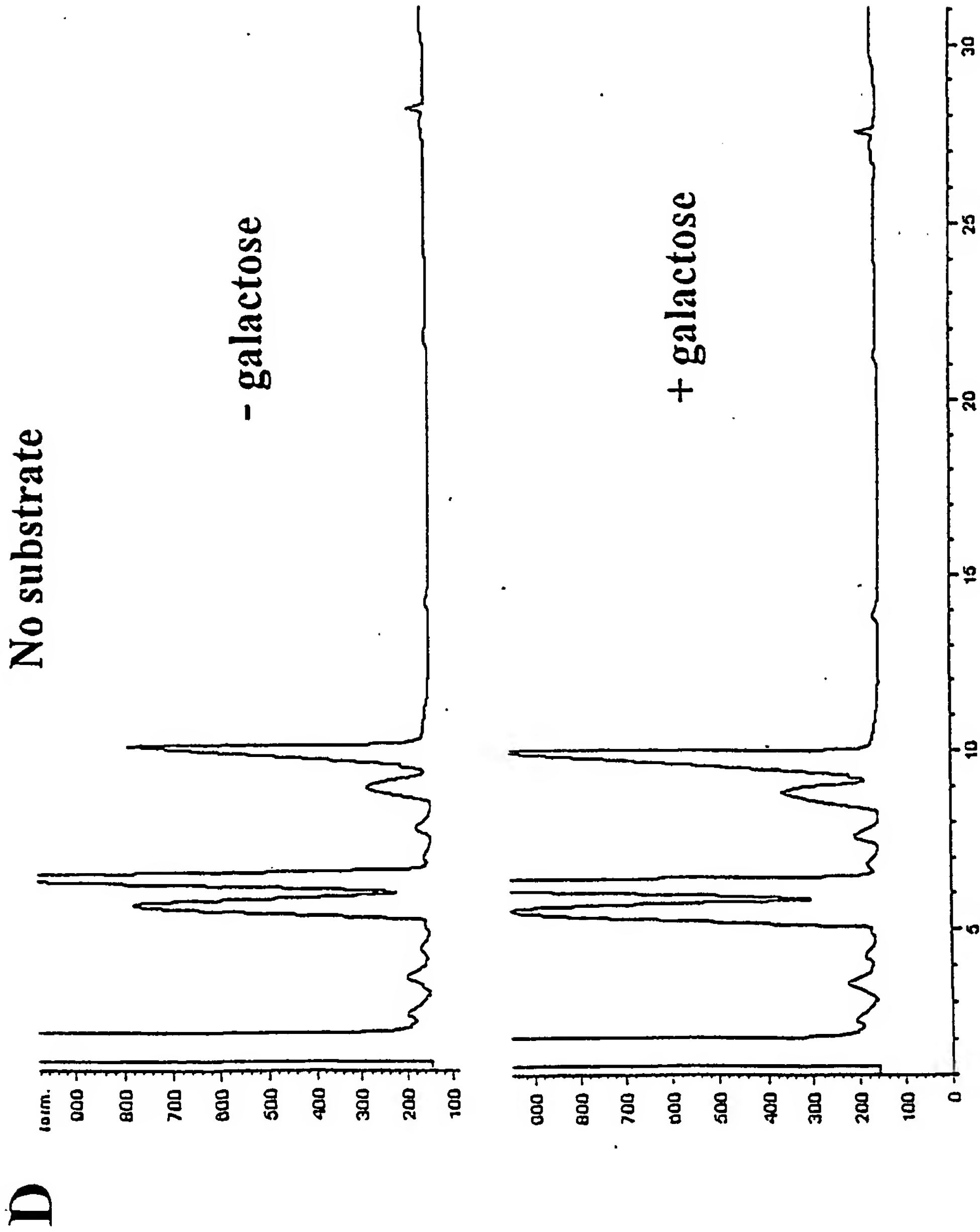




Figure 3 A-D: HPLC chromatograms of fatty acid methyl esters isolated from yeast expressing the Ig\_ASE1 gene product



## SEQUENZPROTOKOLL

&lt;110&gt; University of Bristol

<120> New elongase gene and a process for the production of  
delta-9-polyunsaturated fatty acids

&lt;130&gt; NAE2091/2001

&lt;140&gt; 2001\_2091

&lt;141&gt; 2001-03-21

&lt;160&gt; 2

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 1064

&lt;212&gt; DNA

&lt;213&gt; Isochrysis galbana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (2)..(790)

&lt;400&gt; 1

```

g atg gcc ctc gca aac gac gcg gga gag cgc atc tgg gcg gct gtg acc 49
  Met Ala Leu Ala Asn Asp Ala Gly Glu Arg Ile Trp Ala Ala Val Thr
    1             5             10             15

gac ccg gaa atc ctc att ggc acc ttc tcg tac ttg cta ctc aaa ccg 97
  Asp Pro Glu Ile Leu Ile Gly Thr Phe Ser Tyr Leu Leu Leu Lys Pro
    20             25             30

ctg ctc cgc aat tcc ggg ctg gtg gat gag aag aag ggc gca tac agg 145
  Leu Leu Arg Asn Ser Gly Leu Val Asp Glu Lys Lys Gly Ala Tyr Arg
    35             40             45

acg tcc atg atc tgg tac aac gtt ctg ctg gcg ctc ttc tct gcg ctg 193
  Thr Ser Met Ile Trp Tyr Asn Val Leu Leu Ala Leu Phe Ser Ala Leu
    50             55             60

agc ttc tac gtg acg gcg acc gcc ctc ggc tgg gac tat ggt acg ggc 241
  Ser Phe Tyr Val Thr Ala Thr Ala Leu Gly Trp Asp Tyr Gly Thr Gly
    65             70             75             80

gcg tgg ctg cgc agg caa acc ggc gac aca ccg cag ccg ctc ttc cag 289
  Ala Trp Leu Arg Arg Gln Thr Gly Asp Thr Pro Gln Pro Leu Phe Gln
    85             90             95

tgc ccg tcc ccg gtt tgg gac tcg aag ctc ttc aca tgg acc gcc aag 337
  Cys Pro Ser Pro Val Trp Asp Ser Lys Leu Phe Thr Trp Thr Ala Lys
    100            105            110

gca ttc tat tac tcc aag tac gtg gag tac ctc gac acg gcc tgg ctg 385
  Ala Phe Tyr Tyr Ser Lys Tyr Val Glu Tyr Leu Asp Thr Ala Trp Leu
    115            120            125

gtg ctc aag ggc aag agg gtc tcc ttt ctc cag gcc ttc cac cac ttt 433
  Val Leu Lys Gly Lys Arg Val Ser Phe Leu Gln Ala Phe His His Phe
    130            135            140

ggc gcg ccg tgg gat gtg tac ctc ggc att cgg ctg cac aac gag ggc 481

```

2/3

Gly Ala Pro Trp Asp Val Tyr Leu Gly Ile Arg Leu His Asn Glu Gly  
 145 150 155 160  
 gta tgg atc ttc atg ttt ttc aac tcg ttc att cac acc atc atg tac 529  
 Val Trp Ile Phe Met Phe Phe Asn Ser Phe Ile His Thr Ile Met Tyr  
 165 170 175  
 acc tac tac ggc ctc acc gcc gcc ggg tat aag ttc aag gcc aag ccg 577  
 Thr Tyr Tyr Gly Leu Thr Ala Ala Gly Tyr Lys Phe Lys Ala Lys Pro  
 180 185 190  
 ctc atc acc gcg atg cag atc tgc cag ttc gtg ggc ggc ttc ctg ttg 625  
 Leu Ile Thr Ala Met Gln Ile Cys Gln Phe Val Gly Gly Phe Leu Leu  
 195 200 205  
 gtc tgg gac tac atc aac gtc ccc tgc ttc aac tcg gac aaa ggg aag 673  
 Val Trp Asp Tyr Ile Asn Val Pro Cys Phe Asn Ser Asp Lys Gly Lys  
 210 215 220  
 ttg ttc agc tgg gct ttc aac tat gca tac gtc ggc tcg gtc ttc ttg 721  
 Leu Phe Ser Trp Ala Phe Asn Tyr Ala Tyr Val Gly Ser Val Phe Leu  
 225 230 235 240  
 ctc ttc tgc cac ttt ttc tac cag gac aac ttg gca acg aag aaa tcg 769  
 Leu Phe Cys His Phe Phe Tyr Gln Asp Asn Leu Ala Thr Lys Lys Ser  
 245 250 255  
 gcc aag gcg ggc aag cag ctc taggcctcga gccggctcgc gggttcaagg 820  
 Ala Lys Ala Gly Lys Gln Leu  
 260  
 agggcgacac gggggtggga cgtctgcatg gagatggatt gtggatgtcc ttacgcctta 880  
 ctcatcaatg tcctcccatc tctcccctct agaccttcta ctagccatct agaagggcag 940  
 ctcagagacg gataccgttc cccctcccct tccttttcgt ctttgctttg ccattgtttg 1000  
 tttgtctcta ttttttaaac tattgacgct aacgcgttac gctcgcaaaa aaaaaaaaaa 1060  
 aaaa 1064

<210> 2  
 <211> 263  
 <212> PRT  
 <213> Isochrysis galbana

<400> 2  
 Met Ala Leu Ala Asn Asp Ala Gly Glu Arg Ile Trp Ala Ala Val Thr  
 1 5 10 15  
 Asp Pro Glu Ile Leu Ile Gly Thr Phe Ser Tyr Leu Leu Leu Lys Pro  
 20 25 30  
 Leu Leu Arg Asn Ser Gly Leu Val Asp Glu Lys Lys Gly Ala Tyr Arg  
 35 40 45  
 Thr Ser Met Ile Trp Tyr Asn Val Leu Leu Ala Leu Phe Ser Ala Leu  
 50 55 60  
 Ser Phe Tyr Val Thr Ala Thr Ala Leu Gly Trp Asp Tyr Gly Thr Gly  
 65 70 75 80

3/3

Ala	Trp	Leu	Arg	Arg	Gln	Thr	Gly	Asp	Thr	Pro	Gln	Pro	Leu	Phe	Gln			
				85					90					95				
Cys	Pro	Ser	Pro	Val	Trp	Asp	Ser	Lys	Leu	Phe	Thr	Trp	Thr	Ala	Lys			
			100					105					110					
Ala	Phe	Tyr	Tyr	Ser	Lys	Tyr	Val	Glu	Tyr	Leu	Asp	Thr	Ala	Trp	Leu			
		115					120					125						
Val	Leu	Lys	Gly	Lys	Arg	Val	Ser	Phe	Leu	Gln	Ala	Phe	His	His	Phe			
	130					135					140							
Gly	Ala	Pro	Trp	Asp	Val	Tyr	Leu	Gly	Ile	Arg	Leu	His	Asn	Glu	Gly			
145					150					155				160				
Val	Trp	Ile	Phe	Met	Phe	Phe	Asn	Ser	Phe	Ile	His	Thr	Ile	Met	Tyr			
				165					170					175				
Thr	Tyr	Tyr	Gly	Leu	Thr	Ala	Ala	Gly	Tyr	Lys	Phe	Lys	Ala	Lys	Pro			
			180					185					190					
Leu	Ile	Thr	Ala	Met	Gln	Ile	Cys	Gln	Phe	Val	Gly	Gly	Phe	Leu	Leu			
		195					200					205						
Val	Trp	Asp	Tyr	Ile	Asn	Val	Pro	Cys	Phe	Asn	Ser	Asp	Lys	Gly	Lys			
	210					215					220							
Leu	Phe	Ser	Trp	Ala	Phe	Asn	Tyr	Ala	Tyr	Val	Gly	Ser	Val	Phe	Leu			
225					230					235					240			
Leu	Phe	Cys	His	Phe	Phe	Tyr	Gln	Asp	Asn	Leu	Ala	Thr	Lys	Lys	Ser			
			245						250					255				
Ala	Lys	Ala	Gly	Lys	Gln	Leu												
			260															



(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
3 October 2002 (03.10.2002)

PCT

(10) International Publication Number  
**WO 02/077213 A3**

(51) International Patent Classification<sup>7</sup>: **C12N 15/54**,  
15/82, 9/10, 5/10, C07K 16/40, C07C 57/02, C12P 7/64,  
C12Q 1/68, A01H 5/00

(21) International Application Number: PCT/EP02/03418

(22) International Filing Date: 26 March 2002 (26.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
0107510.0 26 March 2001 (26.03.2001) GB

(71) Applicant (*for all designated States except US*): **UNIVERSITY OF BRISTOL** [GB/GB]; Long Ashton, Bristol BS41 9AF (GB).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **NAPIER, Johnathan, A.** [GB/GB]; 40 The Nursery, Bedminster, Bristol BS3 3EB (GB). **LAZARUS, Colin, M.** [GB/GB]; 119 York Road, Montpelier, Bristol BS6 5QG (GB). **QI, Baoxiu** [GB/GB]; 4 Cumberland House, Norfolk Crescent, Bath BA1 2BG (GB). **LERCHL, Jens** [DE/SE]; Onsjövägen 17, S-26831 Svalöv (SE).

(74) Agent: **DÖRPER, Thomas**; BASF Aktiengesellschaft, 67056 Ludwigshafen (DE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:  
13 March 2003

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: NEW ELONGASE GENE AND PRODUCTION OF  $\Delta^9$ -POLYUNSATURATED FATTY ACIDS

(57) Abstract: The invention relates to a new elongase gene having the sequence SEQ ID NO: 1 or its derivatives, to a gene construct comprising this sequence or its derivatives and to its use. The inventive nucleic acid sequence encodes a polypeptide which elongates  $\alpha$ -linolenic acid ( $C_{18:3}$  d9, 12, 15?) by at least two carbon atoms whereas  $\gamma$ -linolenic acid ( $C_{18:3}$  d6, 9, 12?) is not elongated. The invention additionally relates to vectors or organisms comprising an elongase gene having the sequence SEQ ID NO: 1 or its derivatives. The invention further relates to a process for the production of polyunsaturated fatty acids (=PUFAs) with an organism which comprises the elongase gene and which organism produces high amounts of oils and especially oils with a high content of unsaturated fatty acids. Additionally the invention relates to an oil and/or fatty acid composition with a higher content of polyunsaturated  $C_{20}$  or  $C_{22}$  fatty acids with at least two double bonds and/or to a triacylglycerol composition with a higher content of said polyunsaturated fatty acids.



**WO 02/077213 A3**

## INTERNATIONAL SEARCH REPORT

Inte 1al Application No  
PCT/EP 02/03418

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/54 C12N15/82 C12N9/10 C12N5/10 C07K16/40  
C07C57/02 C12P7/64 C12Q1/68 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K C12P C12Q A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, SEQUENCE SEARCH, MEDLINE, BIOSIS, EMBL

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FERNANDEZ SEVILLA J M ET AL: "Photolimitation and photoinhibition as factors determining optimal dilution rate to produce eicosapentaenoic acid from cultures of the microalga Isochrysis galbana" APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, SPRINGER VERLAG, BERLIN, DE, vol. 50, no. 2, August 1998 (1998-08), pages 199-205, XP002200199 ISSN: 0175-7598	11,12, 20-22
Y	p. 199-200,204, table 1  ---  -/--	2-10, 13-19, 23-25

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

19 September 2002

Date of mailing of the international search report

07/10/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Puonti-Kaerlas, J

## INTERNATIONAL SEARCH REPORT

Inte      nal Application No

PCT/EP 02/03418

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MEDINA A R ET AL: "CONCENTRATION AND PURIFICATION OF STEARIDONIC, EICOSAPENTAENOIC, AND DOCOSAHEXAENOIC ACIDS FROM COD LIVER OIL AND THE MARINE MICROALGA ISOCHRYSIS GALBANA" JOURNAL OF THE AMERICAN OIL CHEMISTS' SOCIETY, AMERICAN OIL CHEMISTS' SOCIETY. CHAMPAIGN, US, vol. 5, no. 72, 1995, pages 575-583, XP001071204 ISSN: 0003-021X p.575-576,582-583, table 2, fig 2,3 ---	20-22
X	WO 00 55330 A (UNIV BRISTOL ;NAPIER JOHNATHAN A (GB)) 21 September 2000 (2000-09-21)	20-22
Y	p. 3:8-6:32;7:25-9:19, claims ---	2-19, 23-25
P,X	QI A B ET AL: "Identification of a cDNA encoding a novel C18-DELTApolyunsaturated fatty acid-specific elongating activity from the docosahexaenoic acid (DHA)-producing microalga, Isochrysis galbana" FEBS LETTERS, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 510, no. 3, 16 January 2002 (2002-01-16), pages 159-165, XP004332633 ISSN: 0014-5793 p.160-164, fig 2 ---	2-25
P,X	WO 02 08401 A (ABBOTT LAB) 31 January 2002 (2002-01-31)	20-22
P,Y	p. 5:15-8:26;28:18-34:11, fig 1, claims ---	2-19, 23-25
P,X	WO 01 59128 A (ZAEHRINGER ULRICH ;BASF AG (DE); HEINZ ERNST (DE); LERCHL JENS (DE)) 16 August 2001 (2001-08-16) abstract, p.3:4-15:13, SEQ ID NO 5, claims ---	2,5-25
Y	WO 00 12720 A (ABBOTT LAB) 9 March 2000 (2000-03-09) cited in the application abstract, p. 1:24-4:32, SEQ ID NO 1, fig 6, example III,V,VI,XI,VV, claims ---	2-25
	--- -/--	

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/03418

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>PARKER-BARNES JENNIFER M ET AL: "Identification and characterization of an enzyme involved in the elongation of n-6 and n-3 polyunsaturated fatty acids" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 97, no. 15, 18 July 2000 (2000-07-18), pages 8284-8289, XP002200200 ISSN: 0027-8424 p. 8288, fig 2</p> <p>---</p>	2-25
Y	<p>BEAUDOIN FREDERIC ET AL: "Heterologous reconstitution in yeast of the polyunsaturated fatty acid biosynthetic pathway" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 97, no. 12, 6 June 2000 (2000-06-06), pages 6421-6426, XP002200201 ISSN: 0027-8424 p. 6421, fig 3</p> <p>---</p>	2-25
Y	<p>DATABASE EMBL 'Online! retrieved from EBI Database accession no. AF268031 XP002213659 the whole document</p> <p>---</p>	2-25
A	<p>US 6 140 486 A (METZ JAMES GEORGE ET AL) 31 October 2000 (2000-10-31)</p> <p>-----</p>	



## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 1, 2-25 (partially)

Present claims 1-25 relate to a compound defined by reference to a desirable characteristic or property, namely to an enzyme with ability to elongate alpha-linolenic acid, but not gamma-linolenic acid.

The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts of claims 2-25 relating to SEQ ID NOs: 1 and 2.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/EP 02/03418

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☒ Claims Nos.: 1, 2-25 (partially)  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  
see FURTHER INFORMATION sheet PCT/ISA/210
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
  
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 02/03418

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
WO 0055330	A	21-09-2000	AU	3308500 A		04-10-2000
			EP	1161542 A1		12-12-2001
			WO	0055330 A1		21-09-2000
			NO	20014542 A		18-09-2001
WO 0208401	A	31-01-2002	AU	7798201 A		05-02-2002
			WO	0208401 A2		31-01-2002
WO 0159128	A	16-08-2001	DE	10005973 A1		16-08-2001
			DE	10023893 A1		22-11-2001
			DE	10063387 A1		12-09-2002
			AU	3924401 A		20-08-2001
			WO	0159128 A2		16-08-2001
WO 0012720	A	09-03-2000	US	6403349 B1		11-06-2002
			AU	5696499 A		21-03-2000
			EP	1108039 A2		20-06-2001
			JP	2002523098 T		30-07-2002
			WO	0012720 A2		09-03-2000
US 6140486	A	31-10-2000	BR	9809946 A		01-08-2000
			EP	1003869 A1		31-05-2000
			JP	2002510205 T		02-04-2002
			WO	9855625 A1		10-12-1998

**THIS PAGE BLANK (USPTO)**